

ADENOSYLCOBALAMIN R = CH₂COHN₂, R' = CH₂CH₂CONH₂

FIG. 1

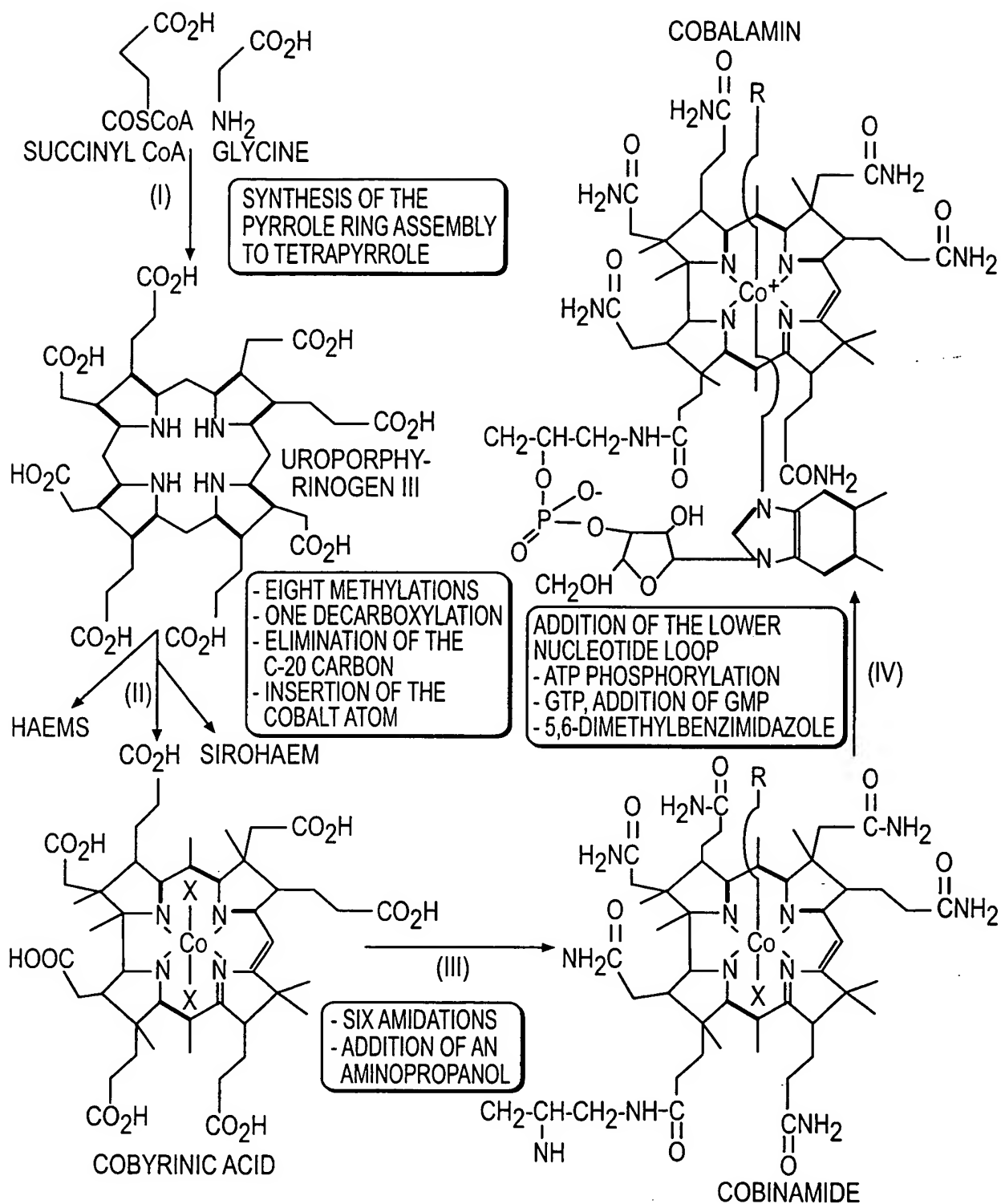


FIG. 2

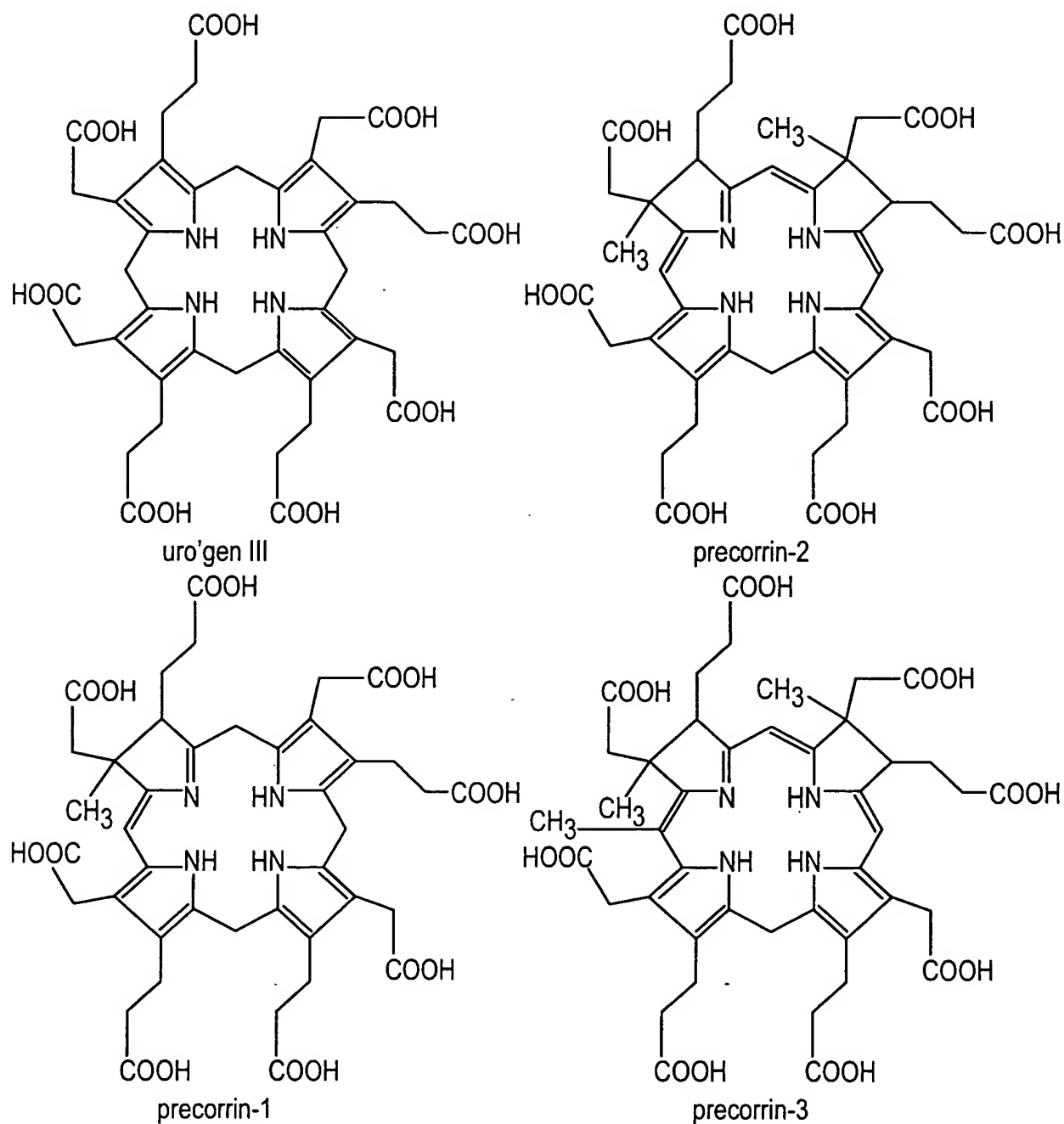
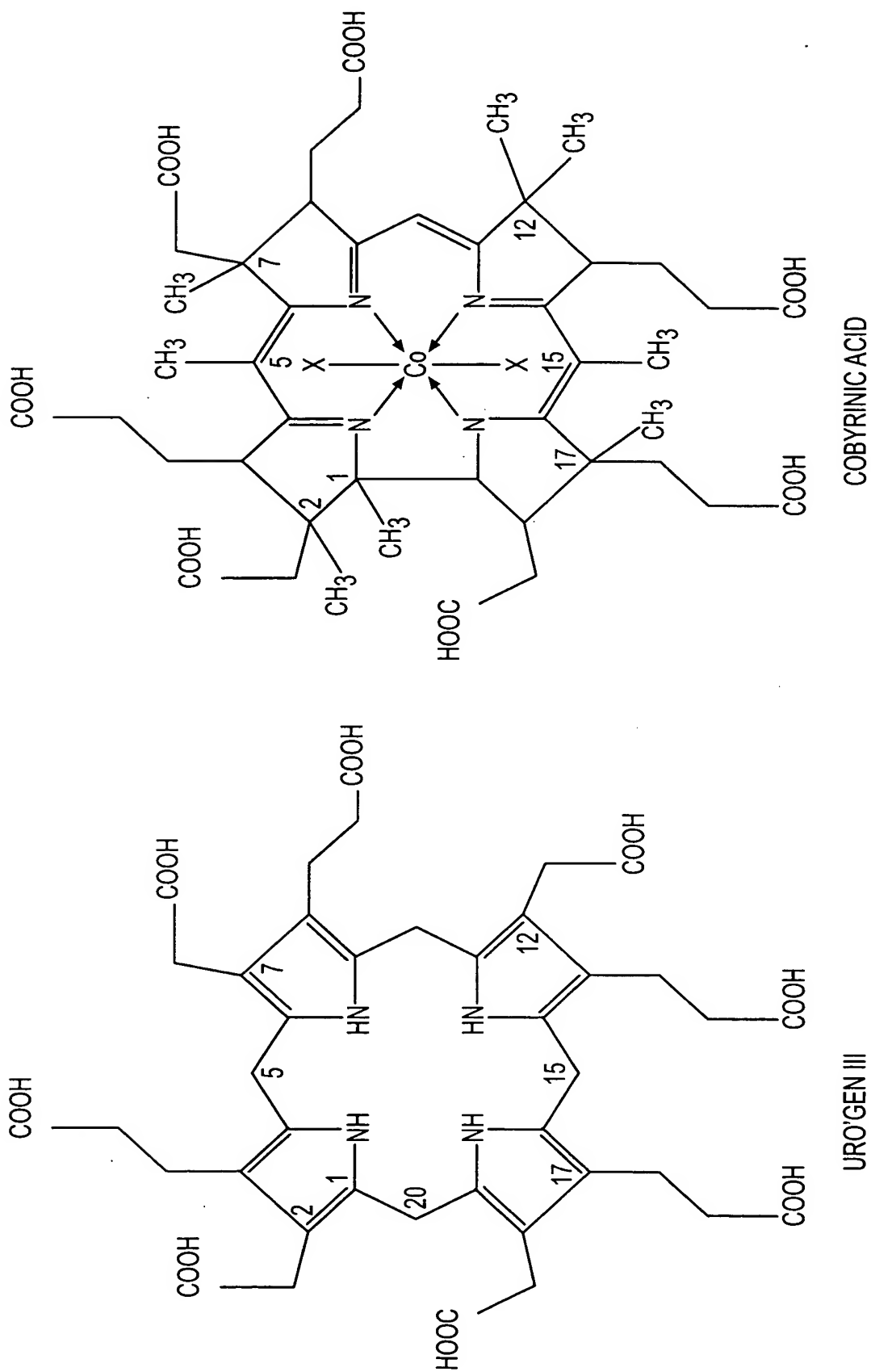


FIG. 3



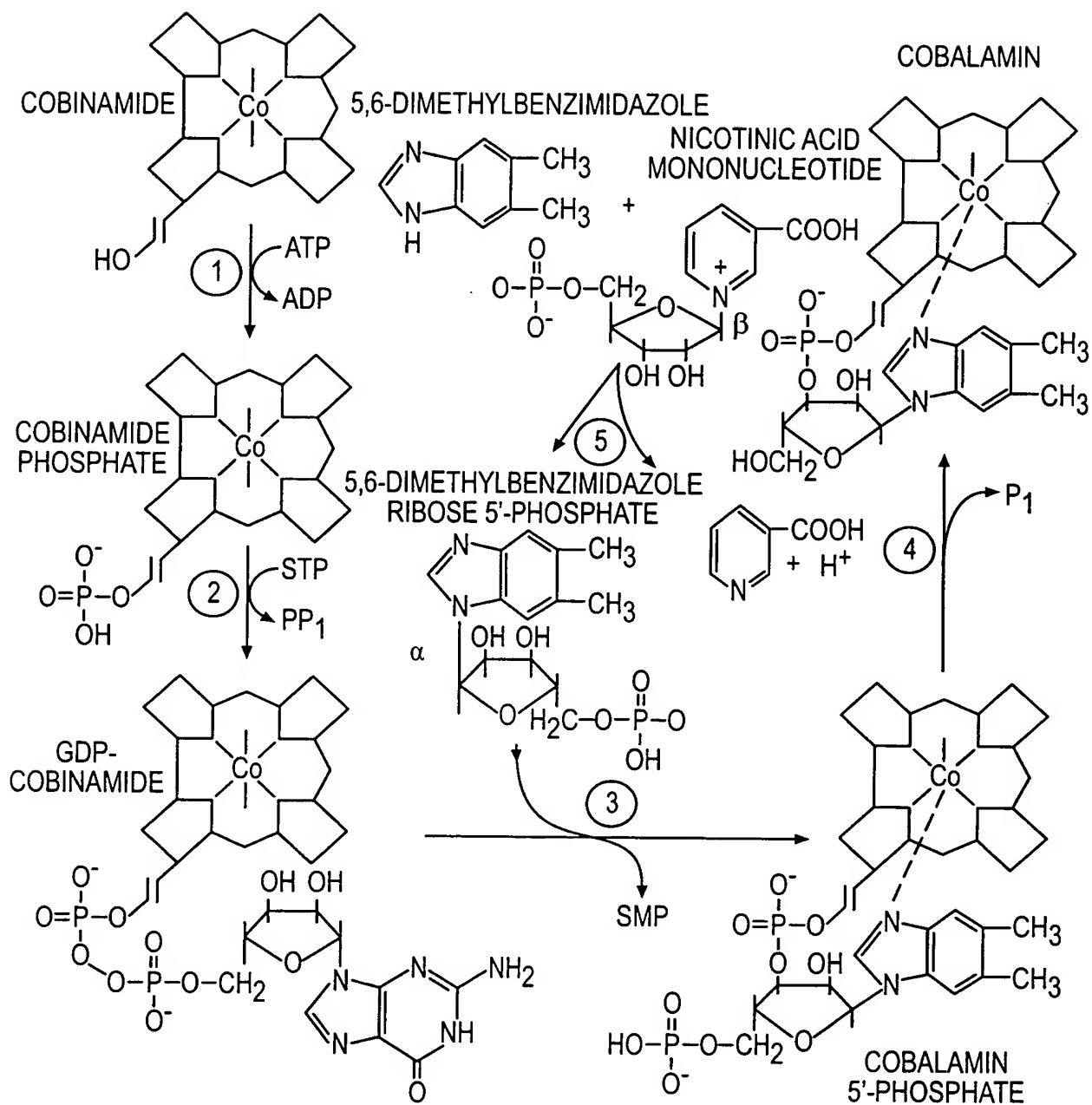


FIG. 5

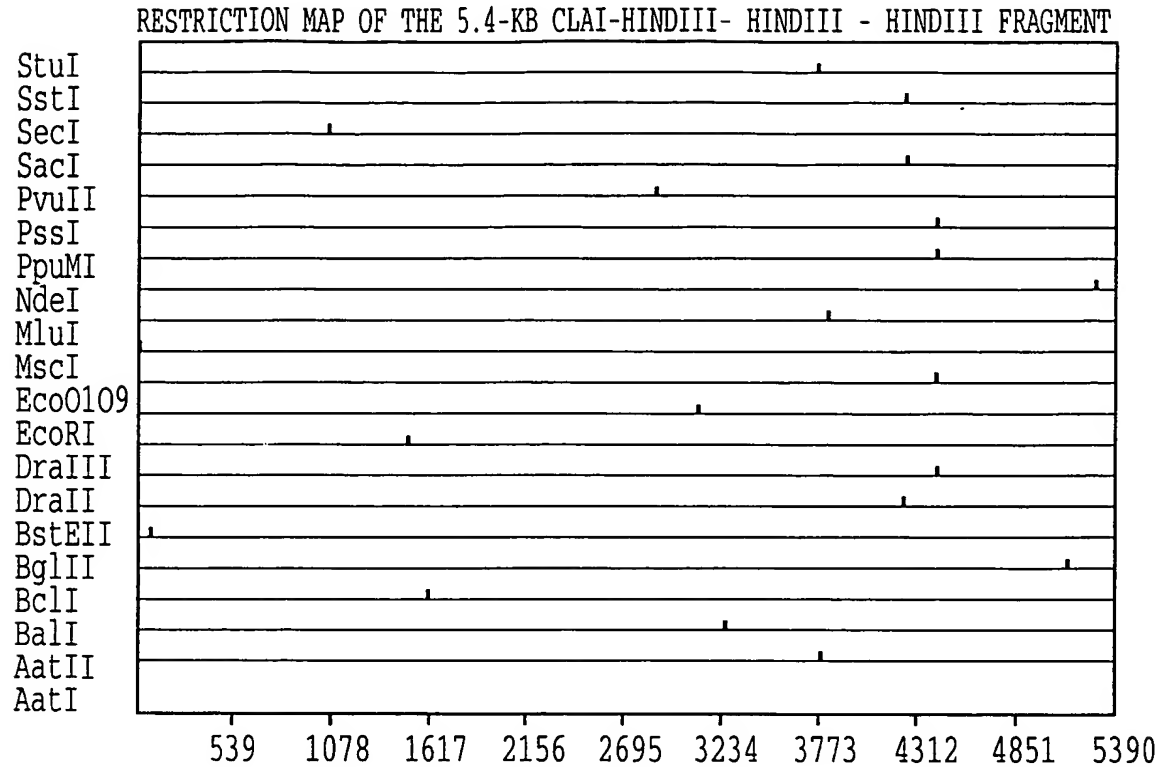


FIG. 6A

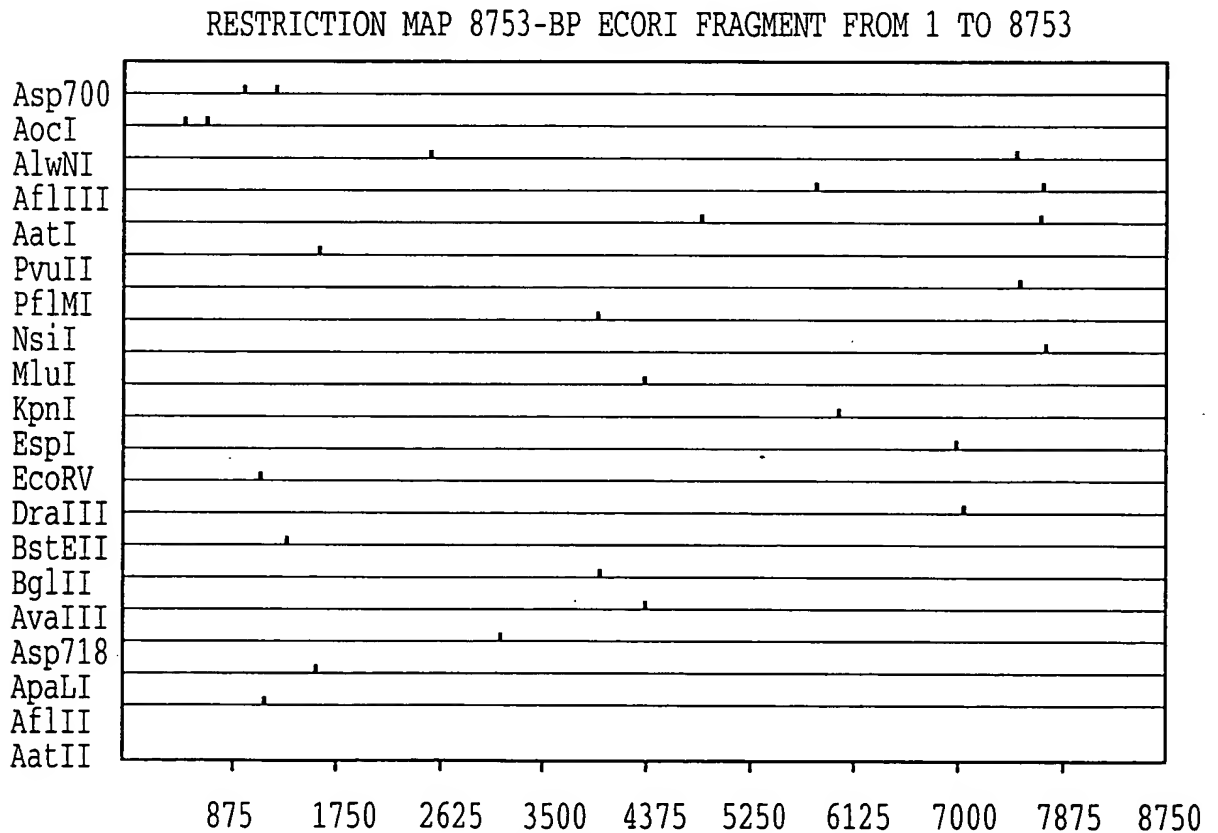


FIG. 6B

RESTRICTION MAP OF THE 4748-BP SalI-SalI-SalI-SalI-SalI-BglI FRAGMENT

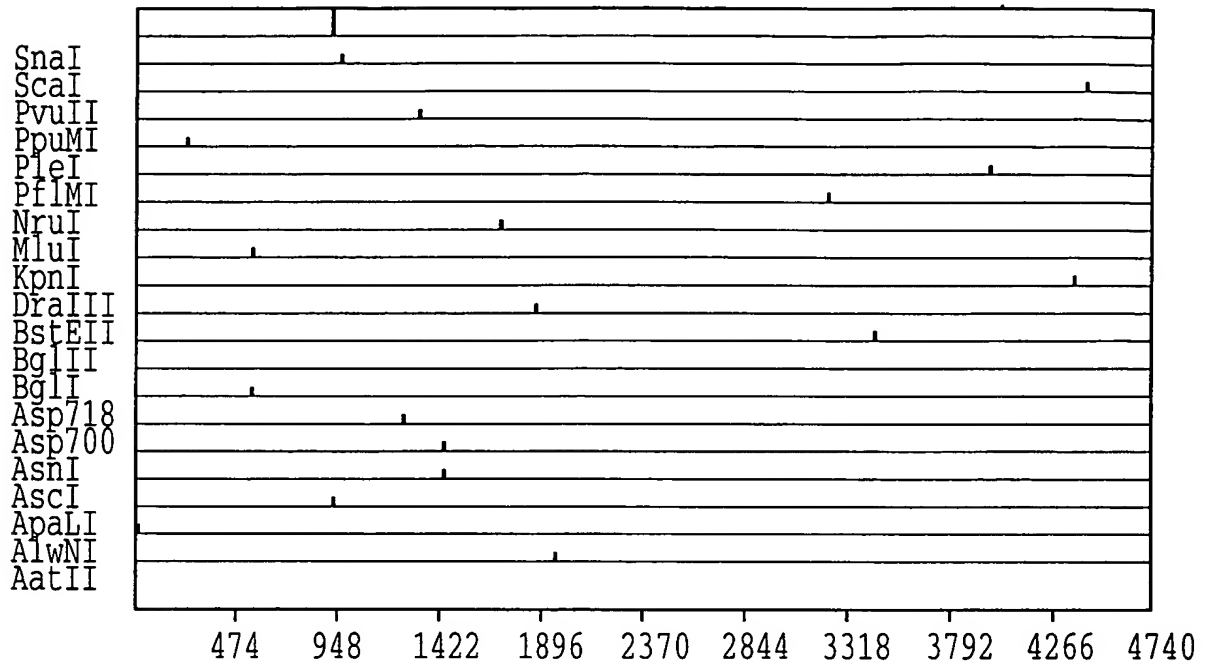


FIG. 6C

RESTRICTION MAP OF THE 3855-BP SatI-SatI-BamHI FRAGMENT

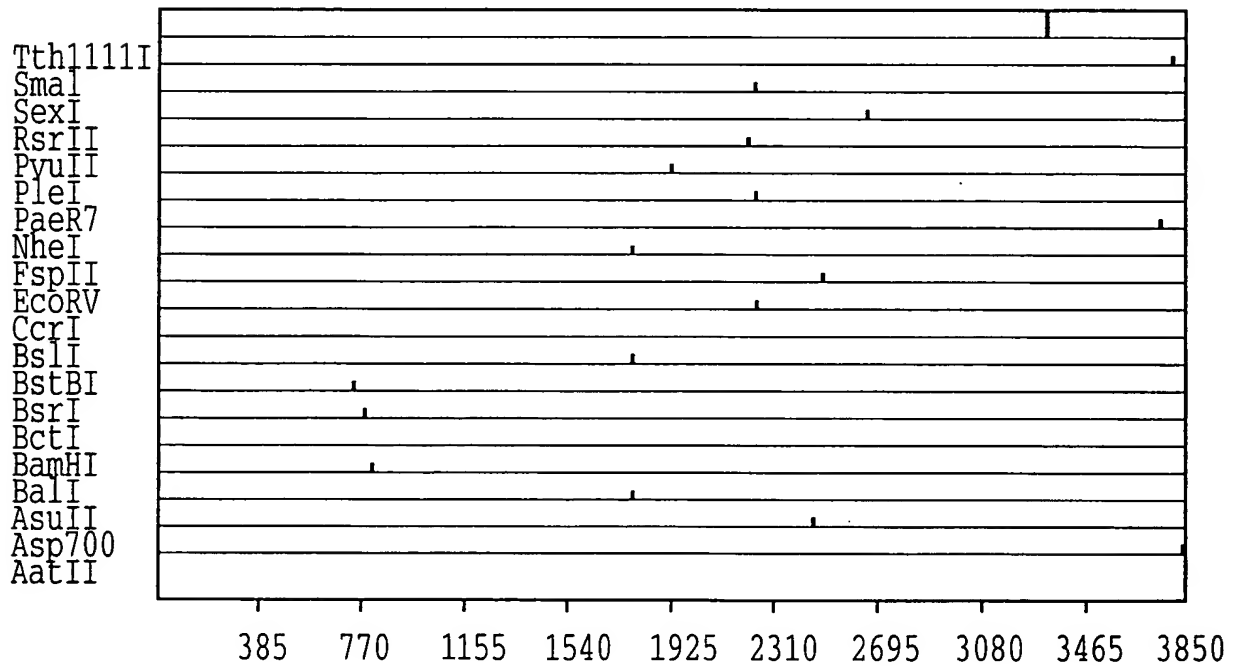


FIG. 6D

10	20	30	40	50	60
GGGCTGCAGG	TCGACTCTAG	AATCGATGAA	GCCTGCGATG	AAGGCGGCCGA	CGAACAGGAA
CCCGACGTCC	AGCTGAGATC	TTAGCTACTT	CGGACGCTAC	TTCCGCCGCT	GCTTGTCTT
70	80	90	100	110	120
GGCGAGCAGG	TGGAAGGCCGA	GATCTTGAC	GGCGGGGACT	CGAGAGGAGA	GCTGTCAGGC
CCGCTCGTCC	ACCTTCCGCT	CTAGAACGTG	CCGCCCTGA	GCTCTCCTCT	CGACAGTCCG
130	140	150	160	170	180
GGGATTTTCC	GCCTTGTGTC	AGAGCCCGGC	GCGATTTGCA	AAGCCTTCTG	TCGCGGTGTT
CCCTAAAAGG	CGGAACACAG	TCTCGGGCCG	CGCTAAACGT	TTCCGAAGAC	AGCGCCACAA
190	200	210	220	230	240
GCTGTCCATG	CAGGTGTCGA	AATTGAAAAA	CCGACAAAGA	TTCAAAGCCT	TGTTCCAGCT
CGACAGGTAC	GTCCACAGCT	TTAACTTTTT	GGCTGTTTCT	AAGTTTCGGA	ACAAGGTCGA
250	260	270	280	290	300
CGCTGTCTTT	CTGGATGGAG	GCGCTCTCGC	CCGCATGGTG	CCGAAGAAGG	GCTGTCCTTG
CGGACAGAAA	GACCTACCTC	CGCGAGAGCG	GGCGTACCAC	GGCTTCTTCC	CGACAGGAAC
310	320	330	340	350	360
CGATACGGTA	GGCGGATGAC	GATCTTCCTC	AAACGCGACA	TGGCGATGGC	GCAATCCGGT
GCTATGCCAT	CCGCCTACTG	CTAGAAGGAG	TTTGCGCTGT	ACCGCTACCG	CGTTAGGCCA
370	380	390	400	410	420
TTGACCGGCC	TTCCGCGCTC	CGGTAAAAAT	GAAGGATATG	CGACGGCGTC	CGCTTTGGCG
AACTGGCCGG	AAGGCGCGAG	GCCATTTTTA	CTTCCTATAC	GCTGCCGCAG	GCGAAACCGC
430	440	450	460	470	480
GA CTGAAAGA	GCGTCCGGTG	CGGCCGACCC	AGTCAGGGGG	GCATCAGCCG	GTGCTGTCCA
CTGACTTTCT	CGCAGGCCAC	GCCGGCTGGG	TCAGTCCCCC	CGTAGTCGGC	CACGACAGGT
490	500	510	520	530	540
GATCGGCCGG	GACGGATCGT	CCCAGCCGGC	GCTTCGTTAA	GGAGAACAAC	GAAGGGAGCC
CTAGCCGGCC	CTGCCTAGCA	GGGTCGGCCG	CGAAGCAATT	CCTCTTGTTG	CTTCCCTCGG
550	560	570	580	590	600
GGCCGCCGAT	GCCATCGGGC	CAACACTCTG	CACAGACGAC	GAAAGCAGGA	GCCGGGCTGG
CCGGCGGCTA	CGGTAGCCCG	GTTGTGAGAC	GTGTCTGCTG	CTTTCGTCCT	CGGCCCCGACC

FIG. 7A

610	620	630	640	650	660
TGCTCGGGCT	CGGCTGCGAG	CGTCGCACGC	CGGCCGAAGA	GGTGATCGCC	CTTGCCGAGC
ACGAGCCCCA	GCCGACGCTC	GCAGCGTGCG	GCCGGCTTCT	CCACTAGCGG	GAACGGCTCG
670	680	690	700	710	720
GTGCGCTTGC	CGATGCCGGT	GTTGCGCCCC	GCGATCTGCG	GCTGGTCGCC	TCGCTCGATG
CACGCGAACG	GCTACGGCCA	CAACGCGGGC	CGCTAGACGC	CGACCAGCGG	AGCGAGCTAC
730	740	750	760	770	780
CTCGCGCCGA	GGAGCCGGCG	ATCCTGGCGG	CCGCTCAGCA	TTTCGCGGTT	CCGGCCGCGT
GAGCGCGGCT	CCTCGGCCGC	TAGGACCGCC	GGCGAGTCGT	AAAGCGCCAA	GGCCGGCGCA
790	800	810	820	830	840
TCTACGATGC	CGCCACGCTC	GAAGCCGAAG	CTTCCCGGCT	CGCCAACCCG	TCCGAGATCG
AGATGCTACG	GCGGTGCGAG	CTTCGGCTTC	GAAGGGCCGA	GCGGTGGGC	AGGCTCTAGC
850	860	870	880	890	900
TCTTTGCCTA	CACGGGTTGT	CATGGCGTTG	CCGAGGGTGC	AGCGCTCGTC	GGCGCCGGTC
AGAAACGGAT	GTGCCCAACA	GTACCGCAAC	GGCTCCCACG	TCGCGAGCAG	CCGCGGCCAG
910	920	930	940	950	960
GCGAAGCCGT	GCTGATTGTG	CAGAAGATCG	TCTCCGCCCA	TGCGACGGCC	GCACTTGCCG
CGCTTCGGCA	CGACTAACAC	GTCTTCTAGC	AGAGGCGGGT	ACGCTGCCGG	CGTGAACGGC
970	980	990	1000	1010	1020
GGCCGGCGAC	CTTGCGCGCC	GAAAAGCGCA	TCCAGGCGGC	GGAGGCTGTC	TGATGCATTC
CCGGCCGCTG	GAACGCGCGG	CTTTTCGCGT	AGGTCCGCCG	CCTCCGACAG	ACTACGTAAG
1030	1040	1050	1060	1070	1080
TTATGTTGTT	GAATTGAATC	AATCTTTTGC	CCGGGGTTTC	TCTCAAGTGG	AATCCGGTTC
AATACAACAA	CTTAACCTAG	TTAGAAAACG	GGCCCCAAAG	AGAGTTCACC	TTAGGCCAAG
1090	1100	1110	1120	1130	1140
TTTAGAGAGC	GCGTCAGGCG	TGCCGTTGGG	TGGCGCCGAA	ATACAGGTGG	GACAGCACGC
AAATCTCTCG	CGCAGTCCGC	ACGGCAACCC	ACCGCGGCTT	TATGTCCACC	CTGTCGTGCG
1150	1160	1170	1180	1190	1200
ATGATCGACG	ACCTCTTTGC	CGGATTGCCG	GCGCTCGAAA	AAGGTTCCGT	CTGGCTGGTC
TACTAGCTGC	TGGAGAAACG	GCCTAACGGC	CGCGAGCTTT	TTCCAAGCCA	GACCGACCAG

FIG. 7B

1210	1220	1230	1240	1250	1260
GGCGCCGGCC	CCGGCGATCC	CGGCCTGTTG	ACGCTGCATG	CGGCCAATGC	GCTGCGCCAG
CCGCGGCCGG	GGCCGCTAGG	GCCGGACAAC	TGCGACGTAC	GCCGGTTACG	CGACGCGGTC
1270	1280	1290	1300	1310	1320
GCGGATGTGA	TCGTGCATGA	TGCGCTGGTC	AACGAGGATT	GCCTGAAGCT	CGCGCGGCCG
CGCCTACACT	AGCACGTACT	ACGCGACCAG	TTGCTCCTAA	CGGACTTCGA	GCGCGCCGGC
1330	1340	1350	1360	1370	1380
GGCGCCGTGC	TGGAGTTTGC	GGGCAAGCGT	GGCGGCAAGC	CGTCGCCGAA	GCAGCGCGAC
CCGCGGCACG	ACCTCAAACG	CCCGTTCGCA	CCGCCGTTCG	GCAGCGGCTT	CGTCGCGCTG
1390	1400	1410	1420	1430	1440
ATCTCGCTTC	GCCTCGTCGA	ACTCGCGCGC	GCCGGCAACC	GGTGCTGCG	CCTCAAAGGC
TAGAGCGAAG	CGGAGCAGCT	TGAGCGCGCG	CGGCCGTTCG	CCCACGACGC	GGAGTTTCCG
1450	1460	1470	1480	1490	1500
GGCGATCCCT	TCGTCTTCGG	TCGCGGTGGC	GAGGAGGCGC	TGACGCTGGT	CGAACACCAG
CCGCTAGGGA	AGCAGAAGCC	AGCGCCACCG	CTCCTCCGCG	ACTGCGACCA	GCTTGTGGTC
1510	1520	1530	1540	1550	1560
GTGCCGTTCC	GAATCGTGCC	CGGCATCACC	GCCGGTATCG	GCGGGCTTGC	CTATGCCGGC
CACGGCAAGG	CTTAGCACGG	GCCGTAGTGG	CGGCCATAGC	CGCCCGAACG	GATACGGCCG
1570	1580	1590	1600	1610	1620
ATTCCCCTGA	CCCATCGCGA	GGTCAACCAC	GCGGTCACCT	TCCTGACTGG	CCATGATTCC
TAAGGGCACT	GGGTAGCGCT	CCAGTTGGTG	CGCCAGTGAA	AGGACTGACC	GGTACTAAGG
1630	1640	1650	1660	1670	1680
TCCGGCCTGG	TGCCGGATCG	CATCAACTGG	CAGGGCATCG	CCAGCGGCTC	GCCTGTCATC
AGGCCGGACC	ACGGCCTAGC	GTAGTTGACC	GTCCCGTAGC	GGTCGCCGAG	CGGACAGTAG
1690	1700	1710	1720	1730	1740
GTCATGTACA	TGGCGATGAA	ACATATCGGC	GCGATCACCG	CCAACCTCAT	TGCCGGCGGC
CAGTACATGT	ACCGCTACTT	TGTATAGCCG	CGCTAGTGGC	GGTTGGAGTA	ACGGCCGCCG
1750	1760	1770	1780	1790	1800
CGCTCGCCGG	ACGAACCGGT	CGCCTTCGTC	TGCAACGCCG	CGACGCCGCA	GCAGGCGGTG
GCGAGCGGCC	TGCTTGGCCA	GCGGAAGCAG	ACGTTGCGGC	GCTGCGGCGT	CGTCCGCCAC

FIG. 7C

1810	1820	1830	1840	1850	1860
CTGGAAACGA	CGCTTGCGCG	TGCAGAGGCC	GATGTTGCGG	CGGCAGGGCT	GGAGCCGCCG
GACCTTTGCT	GCGAACGCGC	ACGTCTCCGG	CTACAACGCC	GCCGTCCCGA	CCTCGGCGGC
1870	1880	1890	1900	1910	1920
GCGATCGTCG	TCGTGCGCGA	GGTGGTGCGG	CTGCGCGCAG	CGCTCGACTG	GATCGGCGCG
CGCTAGCAGC	AGCAGCCGCT	CCACCACGCC	GACGCGCGTC	GCGAGCTGAC	CTAGCCGCGC
1930	1940	1950	1960	1970	1980
CTGGACGGGC	GCAAGCTTGC	CGCCGACCCG	TTCGCCAATC	GCATTCTCAG	GAACCCGGCA
GACCTGCCCC	CGTTCGAACG	GCGGCTGGGC	AAGCGGTTAG	CGTAAGAGTC	CTTGGGCCGT
1990	2000	2010	2020	2030	2040
TGAGCGGATT	GCTGATTGCC	GCACCCGCGT	CCGGCTCCGG	CAAGACGACG	GTGACGCTCG
ACTCGCCTAA	CGACTAACGG	CGTGGGCGCA	GGCCGAGGCC	GTTCTGCTGC	CACTGCGAGC
2050	2060	2070	2080	2090	2100
GGCTGATGCG	CGCCCTGAAG	AGGCGCGGCG	TGGCGATCGC	GCCCGGCAAG	GCGGGGCCCG
CCGACTACGC	GCGGGACTTC	TCCGCGCCGC	ACCGCTAGCG	CGGGCCGTTC	CGCCCCGGCC
2110	2120	2130	2140	2150	2160
ACTATATCGA	TCCCGCTTTC	CACGCGGCAG	CGACCGGCGA	GCCCTGCTTC	AACTACGACC
TGATATAGCT	AGGGCGAAAG	GTGCGCCGTC	GCTGGCCGCT	CGGGACGAAG	TTGATGCTGG
2170	2180	2190	2200	2210	2220
CCTGGGCGAT	GCGCCCGGAA	CTGCTGCTTG	CCAATGCGTC	GCATGTGGCC	TCCGGCGGGC
GGACCCGCTA	CGCGGGCCTT	GACGACGAAC	GGTTACGCAG	CGTACACCGG	AGGCCGCCCG
2230	2240	2250	2260	2270	2280
GCACATTGAT	CGTCGAGGCG	ATGATGGGAC	TGCATGACGG	TGCTGCCGAC	GGCTCGGGAA
CGTGTAAGTA	GCAGCTCCGC	TACTACCCTG	ACGTACTGCC	ACGACGGCTG	CCGAGCCCTT
2290	2300	2310	2320	2330	2340
CGCCAGCGGA	CCTCGCCGCG	ACGCTGAACC	TTGCGGTCAT	TCTGGTGGTC	GATTGCGCCC
GCGGTCGCCT	GGAGCGGCGC	TGCGACTTGG	AACGCCAGTA	AGACCACCAG	CTAACGCGGG
2350	2360	2370	2380	2390	2400
GCATGTCCCA	GTCGGTTGCC	GCCCTCGTGC	GCGGCTATGC	GGATCATCGC	GACGATATCC
CGTACAGGGT	CAGCCAACGG	CGGGAGCACG	CGCCGATACG	CCTAGTAGCG	CTGCTATAGG

FIG. 7D

2410	2420	2430	2440	2450	2460
GGGTGGTTGG	CGTCATCCTC	AACAAGGTCG	GCAGCGATCG	GCATGAAATG	ATGCTGCGCG
CCCACCAACC	GCAGTAGGAG	TTGTTCCAGC	CGTCGCTAGC	CGTACTTTAC	TACGACGCGC
2470	2480	2490	2500	2510	2520
ATGCGCTCGG	CAAGGTGCGC	ATGCCTGTCT	TCGGCGTGCT	CCGGCAGGAC	AGCGCATTGC
TACGCGAGCC	GTTCCACGCG	TACGGACAGA	AGCCGCACGA	GGCCGTCCTG	TCGCGTAACG
2530	2540	2550	2560	2570	2580
AACTGCCGGA	GCGCCATCTC	GGGCTCGTGC	AGGCGGGCGA	ACACTCAGCG	CTTGAGGGCT
TTGACGGCCT	CGCGGTAGAG	CCCAGACAG	TCCGCCCGCT	TGTGAGTCGC	GAAC TCCCGA
2590	2600	2610	2620	2630	2640
TCATCGAGGC	GGCGGCCGCG	CGGGTCGAGG	CTGCCTGCGA	TCTCGACGCC	ATCCGCCTGA
AGTAGCTCCG	CCGCCGGGCG	GCCCAGCTCC	GACGGACGCT	AGAGCTGCGG	TAGGCGGACT
2650	2660	2670	2680	2690	2700
TCGCGACGAT	TTTCCCGCAG	GTGCCC GCGG	CGGCCGATGC	CGAGCGTTTG	CGGCCGCTCG
AGCGCTGCTA	AAAGGGCGTC	CACGGGCGCC	GCCGGCTACG	GCTCGCAAAC	GCCGGCGAGC
2710	2720	2730	2740	2750	2760
GTCAGCGCAT	CGCGGTGCGG	CGCGATATCG	CCTTTGCCTT	CTGCTACGAG	CACCTGCTTT
CAGTCGCGTA	GCGCCAGCGC	GCGCTATAGC	GGAAACGGAA	GACGATGCTC	GTGGACGAAA
2770	2780	2790	2800	2810	2820
ACGGCTGGCG	GCAAGGCGGC	GCGGAGATTT	CCTTCTTCTC	GCCGCTCGCC	GACGAGGGGC
TGCCGACCGC	CGTTCCGCCG	CGCCTCTAAA	GGAAGAAGAG	CGGCGAGCGG	CTGCTCCCCG
2830	2840	2850	2860	2870	2880
CGGATGCGGC	AGCCGATGCC	GTCTATCTTC	CGGGGGGTTA	TCCGGAGCTG	CATGCGGGGC
GCCTACGCCG	TCGGCTACGG	CAGATAGAAG	GCCCCCAAT	AGGCCTCGAC	GTACGCCCCG
2890	2900	2910	2920	2930	2940
AGCTGAGCGC	CGCCGCCCGA	TTCCGTTCCG	GCATGCATTC	CGCGGCGGAA	CGCGGCGCCC
TCGACTCGCG	GCGGCGGGCT	AAGGCAAGGC	CGTACGTAAG	GCGCCGCCCT	GCGCCGCGGG
2950	2960	2970	2980	2990	3000
GCATCTTCGG	CGAGTGCGGC	GGCTATATGG	TGCTCGGCGA	AGGGCTTGTC	GCTGCCGATG
CGTAGAAGCC	GCTCACGCCG	CCGATATACC	ACGAGCCGCT	TCCCGAACAG	CGACGGCTAC

FIG. 7E

3010	3020	3030	3040	3050	3060
GCACACGCTA	CGACATGCTC	GGCCTGCTGC	CGCTCGTAAC	CAGTTTTGCC	GAGCGCAGGC
CGTGTGCGAT	GCTGTACGAG	CCGGACGACG	GCGAGCATTG	GTCAAAACGG	CTCGCGTCCG
3070	3080	3090	3100	3110	3120
GGCACCTCGG	CTATCGCCGC	GTCGTGCCTG	TCGACAACGC	CTTCTTCGAT	GGACCCATGA
CCGTGGAGCC	GATAGCGGCG	CAGCACGGAC	AGCTGTTGCG	GAAGAAGCTA	CCTGGGTACT
3130	3140	3150	3160	3170	3180
CGGCGCACGA	ATTCCACTAT	GCGACCATCG	TCGCCGAAGG	GGCGGCCGAT	CGGCTGTTTG
GCCGCGTGCT	TAAGGTGATA	CGCTGGTAGC	AGCGGCTTCC	CCGCCGGCTA	GCCGACAAAC
3190	3200	3210	3220	3230	3240
CGGTCAGCGA	CGCCGCCGGC	GAGGATCTCG	GCCAGGCGGG	CCTCCGGCGC	GGCCCTGTGC
GCCAGTCGCT	GCGGCGGCCG	CTCCTAGAGC	CGGTCCGCCC	GGAGGCCGCG	CCGGGACAGC
3250	3260	3270	3280	3290	3300
CCGGTTCCTT	CATGCATCTG	ATCGACGTGC	CAGGTGCTGC	ATGAGCGCAC	CGATCGTTCA
GGCCAAGGAA	GTACGTAGAC	TAGCTGCAGC	GTCCACGACG	TACTCGCGTG	GCTAGCAAGT
3310	3320	3330	3340	3350	3360
TGGTGGCGGC	ATCACCGAGG	CCGCAGCGCG	CTATGGCGGC	CGGCCTGAAG	ACTGGCTCGA
ACCACCGCCG	TAGTGGCTCC	GGCGTCGCGC	GATACCGCCG	GCCGGACTTC	TGACCGAGCT
3370	3380	3390	3400	3410	3420
TCTGTGACCG	GGCATCAATC	CATGCCCCGT	CGCCTTGCCC	GCGGTCCCTG	AGCGCGCCTG
AGACAGCTGG	CCGTAGTTAG	GTACGGGGCA	GCGGAACGGG	CGCCAGGGAC	TCGCGCGGAC
3430	3440	3450	3460	3470	3480
GCACCGGCTG	CCGGATCGGC	AGACGGTAGA	TGATGCGCGG	AGCGCCGCCG	CCGACTACTA
CGTGGCCGAC	GGCCTAGCCG	TCTGCCATCT	ACTACGCGCC	TCGCGGCGGC	GGCTGATGAT
3490	3500	3510	3520	3530	3540
CCGCACCAAC	GGCGTGCTGC	CTTTGCCGGT	GCCGGGCACC	CAGTCGGTGA	TCCAGCTCCT
GGCGTG GTTG	CCGCACGACG	GAAACGGCCA	CGGCCC GTGG	GTCAGCCACT	AGGTCGAGGA
3550	3560	3570	3580	3590	3600
GCCACGTCTT	GCTCCGGCCA	ACAGGCACGT	CGCGATTTTC	GGGCCGACCT	ATGGCGAGTA
CGGTGCAGAA	CGAGGCCGGT	TGTCCGTGCA	GCGCTAAAAG	CCCGGCTGGA	TACCGCTCAT

FIG. 7F

3610	3620	3630	3640	3650	3660
TGCCCCGCGTG	CTTGAAGCGG	CCGGCTTTTGC	TGTCGATCGC	GTCGCGGATG	CCGACGCGCT
ACGGGCGCAC	GAAC TTCGCC	GGCCGAAACG	ACAGCTAGCG	CAGCGCCTAC	GGCTGCGCGA
3670	3680	3690	3700	3710	3720
CACGGCCGAA	CATGGGCTTG	TCATCGTTCG	CAACCCCAAC	AACCCGACCG	GCCGCGCCTT
GTGCCGGCTT	GTACCCGAAC	AGTAGCAGCA	GTTGGGGTTG	TTGGGCTGGC	CGGCGCGGAA
3730	3740	3750	3760	3770	3780
GGCGCCGGCG	GAGCTTCTGG	CGATCGCCGC	AAGGCAGAAG	GCGAGCGGCG	GACTGCTGCT
CCGCGGCCGC	CTCGAAGACC	GCTAGCGGCG	TTCCGTCTTC	CGCTCGCCGC	CTGACGACGA
3790	3800	3810	3820	3830	3840
GGTCGATGAG	GCCTTCGGCG	ATCTTGAGCC	GCAACTGAGT	GTCGCTGGTC	ACGCGTCAGG
CCAGCTACTC	CGGAAGCCGC	TAGA ACTCGG	CGTTGACTCA	CAGCGACCAG	TGCGCAGTCC
3850	3860	3870	3880	3890	3900
GCAAGGCAAC	CTCATCGTCT	TCCGCTCCTT	CGGCAAGTTC	TTGGGCCTTG	CGGGCCTGCG
CGTTCCGTTG	GAGTAGCAGA	AGGCGAGGAA	GCCGTTCAAG	AAGCCGGAAC	GCCCCGACGC
3910	3920	3930	3940	3950	3960
CCTCGGCTTC	GTCGTTGCGA	CCGAGCCAGT	GCTTG CATCC	TTTGCCGATT	GGCTCGGTCC
GGAGCCGAAG	CAGCAACGCT	GGCTCGGTCA	CGAACGTAGG	AAACGGCTAA	CCGAGCCAGG
3970	3980	3990	4000	4010	4020
CTGGGCTGTC	TCCGGCCCCG	CGTTGACGAT	CTCGAAAGCG	CTGATGCAGG	GCGATACGAA
GACCCGACAG	AGGCCGGGCC	GCAACTGCTA	GAGCTTTTCG	GACTACGTCC	CGCTATGCTT
4030	4040	4050	4060	4070	4080
GGCGATCGCG	GCGGGCATCC	TCGAGCGTCG	CGCCGGCCTC	GATGCGGCTC	TCGATGGGGC
CCGCTAGCGC	CGCCCGTAGG	AGCTCGCAGC	GCGGCCGGAG	CTACGCCGAG	AGCTACCCCG
4090	4100	4110	4120	4130	4140
AGGGCTCAAC	CGTATCGGCG	GCACGGGGCT	ATTCTGCTG	GTCGAGCATC	CCAGGGCAGC
TCCCGAGTTG	GCATAGCCGC	CGTGCCCCGA	TAAGCACGAC	CAGCTCGTAG	GGTCCCGTCG
4150	4160	4170	4180	4190	4200
TCTGCTGCAG	GAGCGGCTCT	GCGAGGCCCA	TATTCTCACG	CGCAAGTTCG	ACTATGCCCC
AGACGACGTC	CTCGCCGAGA	CGTCCGGGT	ATAAGAGTGC	GCGTTCAAGC	TGATACGGGG

FIG. 7G

4210	4220	4230	4240	4250	4260
GACCTGGCTC	AGGGTCGGTC	TTGCGCCTGA	CGCGGCTGGT	GACCGACGGC	TGGCGGACGC
CTGGACCGAG	TCCCAGCCAG	AACGCGGACT	GCGCCGACCA	CTGGCTGCCG	ACCGCCTGCG
4270	4280	4290	4300	4310	4320
GCTTGCCCCG	ATGGAGCTCT	GAGGTGTCGG	AGACGATCCT	GCTCATTCTC	GCGCTGGCGC
CGAACGGGCG	TACCTCGAGA	CTCCACAGCC	TCTGCTAGGA	CGAGTAAGAG	GCGGACCGCG
4330	4340	4350	4360	4370	4380
TGGTGATCGA	CCGCGTTGTC	GGCGATCCGG	ACTGGCTCTG	GGCGCGCGTG	CCGCATCCGG
ACCACTAGCT	GGCGCAACAG	CCGCTAGGCC	TGACCGAGAC	CCGCGCGCAC	GGCGTAGGCC
4390	4400	4410	4420	4430	4440
TCGTGTTTTT	CGGCAAGGCC	ATCGGCTTTT	TCGACGCGCG	GCTGAACCGG	GAGGACCTCG
AGCACAAAAA	GCCGTTCCGG	TAGCCGAAAA	AGCTGCGCGC	CGACTTGGCC	CTCCTGGAGC
4450	4460	4470	4480	4490	4500
AGGATAGCGC	GCGCAAATTT	CGTGGCGTCG	TCGCGATCCT	TTTGTGCTT	GGCATCAGCG
TCCTATCGCG	CGCGTTTAAA	GCACCGCAGC	AGCGCTAGGA	AAACAACGAA	CCGTAGTCGC
4510	4520	4530	4540	4550	4560
CCTGGTTCGG	CCATCTGCTG	CATCGCCTGT	TCGCCGTCTT	CGGACCGCTC	GGCTTTCTGC
GGACCAAGCC	GGTAGACGAC	GTAGCGGACA	AGCGGCAGGA	GCCTGGCGAG	CCGAAAGACG
4570	4580	4590	4600	4610	4620
TCGAGGCGGT	TCTGGTCGCG	GTCTTCCTGG	CACAGAAGAG	CCTCGCCGAT	CACGTGCGTC
AGCTCCGCCA	AGACCAGCGC	CAGAAGGACC	GTGTCTTCTC	GGAGCGGCTA	GTGCACGCAG
4630	4640	4650	4660	4670	4680
GCGTGGCCGG	GGGCTTGCGA	CAGGGCGGGC	TGGAAGGCGG	GCGTGCCGCC	GTGTCGATGA
CGCACCGGCC	CCCGAACGCT	GTCCCGCCCG	ACCTTCCGCC	CGCACGGCGG	CACAGCTACT
4690	4700	4710	4720	4730	4740
TCGTTGGTCG	CGATCCAAAG	ACGCTCGACG	AGCCGGCGGT	CTGCCGTGCC	GCGATCGAAA
AGCAACCAGC	GCTAGGTTTC	TGCGAGCTGC	TCGGCCGCCA	GACGGCACGG	CGCTAGCTTT
4750	4760	4770	4780	4790	4800
GCCTTGCCGA	GAATTTCTCC	GACGGCGTCG	TGGCGCCGGC	CTTCTGGTAC	GCGGTTGCCG
CGGAACGGCT	CTTAAAGAGG	CTGCCGCAGC	ACCGCGGCCG	GAAGACCATG	CGCCAACGGC

FIG. 7H

4810	4820	4830	4840	4850	4860
GCCTGCCGGG	GCTTCTTGCC	TACAAGATGC	TGAACACCGC	CGATTTCGATG	ATCGGCCACA
CGGACGGCCC	CGAAGAACGG	ATGTTCTACG	ACTTGTGGCG	GCTAAGCTAC	TAGCCGGTGT
4870	4880	4890	4900	4910	4920
AGTCGCCGAA	ATATCTGCAC	TTCGGCTGGG	CCTCGGCCCC	ACTCGACGAT	CTCGCCAACC
TCAGCGGCTT	TATAGACGTG	AAGCCGACCC	GGAGCCGGGC	TGAGCTGCTA	GAGCGGTG
4930	4940	4950	4960	4970	4980
TGCCGGCAGC	GAGGCTCTCG	ATCCTTTTGA	TCTCAGCCGG	TGCGCTGATC	CATCGTGCGC
ACGGCCGTCTG	CTCCGAGAGC	TAGGAAACT	AGAGTCGGCC	ACGCGACTAG	GTAGCACC
4990	5000	5010	5020	5030	5040
CCAGCGCCGC	CAAGGATGCG	CTGACCGTGG	CCCTTCGCGA	CCATGGCCTG	CACCGCTCGC
GGTCGCGGCG	GTTCCCTACG	GACTGGCACC	GGGAAGCGCT	GGTACCGGAC	GTGGCGAGCG
5050	5060	5070	5080	5090	5100
CGAACTCCGG	CTGGCCGGAA	GCGGCCATGG	CCGGCGCGCT	CGATCTGCAG	CTTGCCGGTC
GCTTGAGGCC	GACCGGCCTT	CGCCGGTACC	GGCCGCGCGA	GCTAGACGTC	GAACGGCCAG
5110	5120	5130	5140	5150	5160
CGCGGATCTA	TGGCGGCGTC	AAGGTCAGCG	AACCTATGAT	CAACGGTCCG	GGCCGAGCGG
GCGCCTAGAT	ACCGCCGCAG	TTCCAGTCGC	TTGGATACTA	GTTGCCAGGC	CCGGCTCGCC
5170	5180	5190	5200	5210	5220
TTGCAACAAG	CGAAGACATC	GACGCCGGTA	TTGCTGTATT	TTATGGCGCC	TGTACGGTCA
AACGTTGTTC	GCTTCTGTAG	CTGCGGCCAT	AACGACATAA	AATACCGCGG	ACATGCCAGT
5230	5240	5250	5260	5270	5280
TGGCCGGGTT	TGTTCTTGCA	ATCGCAATGA	TTTGATCGCG	GAAGTTGACC	TTCGCATTAA
ACCGGCCCAA	ACAAGAACGT	TAGCGTTACT	AAACTAGCGC	CTTCAACTGG	AAGCGTAATT
5290	5300	5310	5320	5330	5340
GACTCTGCTT	TCCATATGTA	TTAAGATCGT	ATCATATTTCG	ATCAGTTATT	CTCCTGGAAC
CTGAGACGAA	AGGTATACAT	AATTCTAGCA	TAGTATAAGC	TAGTCAATAA	GAGGACCTTG
5350	5360	5370	5380	5390	5400
GTTTGGTTCC	ACCGGTACGT	GTTCGTCTTC	CCGGAGAGAG	AAGCATGCGC	AAAAGCTT
CAAACCAAGG	TGGCCATGCA	CAAGCAGAAG	GGCCTCTCTC	TTCGTACGCG	TTTTCGAA

FIG. 71

10	20	30	40	50	60
GAATTCGCCA	GCGCCTACAT	GGCTGACCTC	AAGCAGTTCC	TCGTGGCCCA	GAAGAACGAG
CTTAAGCGGT	CGCGGATGTA	CCGACTGGAG	TTCGTCAAGG	AGCACCGGGT	CTTCTTGCTC
70	80	90	100	110	120
GGCCGGCAGA	TTTTCCCTCG	CGGGCCTGAG	TATTTTCGCG	CGCTCGACCT	GACGCCGCTC
CCGGCCGTCT	AAAAGGGAGC	GCCCCGACTC	ATAAAAGCGC	GCGAGCTGGA	CTGCGGCGAG
130	140	150	160	170	180
GACAAGGTGC	GCGTGGTCAT	TCTCGGCCAG	GATCCCTATC	ACGGTGACGG	CCAGGCGCAT
CTGTTCCACG	CGCACCAGTA	AGAGCCGGTC	CTAGGGATAG	TGCCACTGCC	GGTCCGCGTA
190	200	210	220	230	240
GGGCTCTGCT	TCAGCGTTCG	CCCCGGTGTC	CGGACGCCGC	CGTCGCTGGT	CAACATCTAC
CCCAGACGA	AGTCGCAAGC	GGGGCCACAG	GCCTGCGGCG	GCAGCGACCA	GTTGTAGATG
250	260	270	280	290	300
AAGGAAGTGA	ATACCGATCT	CGGTATTCCG	CCGGCGCGTC	ACGGTTTTTCT	CGAAAGCTGG
TTCCTTGACT	TATGGCTAGA	GCCATAAGGC	GGCCGCGCAG	TGCCAAAAGA	GCTTTCGACC
310	320	330	340	350	360
GCAAGGCAGG	GCGTGCTGCT	TTTGAACAGC	GTGCTGACGG	TAGAGCGCGG	GAACGTGCGT
CGTTCCGTCC	CGCACGACGA	AAACTTGTCG	CACGACTGCC	ATCTCGCGCC	CTTGACGCA
370	380	390	400	410	420
CACACCAGGG	TCACGGTTGG	GAAAAGTTCA	CGGATGCGAT	CATCCGTGCG	GTCAACGAGG
GTGTGGTCCC	AGTGCCAACC	CTTTTCAAGT	GCCTACGCTA	GTAGGCACGC	CAGTTGCTCC
430	440	450	460	470	480
CCGAGCATCC	CGTCGTCTTC	ATGCTTTGGG	GCTCCTATGC	GCAGAAGAAG	GCGGCCTTCG
GGCTCGTAGG	GCAGCAGAAG	TACGAAACCC	CGAGGATACG	CGTCTTCTTC	CGCCGGAAGC
490	500	510	520	530	540
TCGACCGCTC	GCGCCATCTT	GTCCTGAGGG	CACCACATCC	GTCGCCGCTC	TCAGCCCATT
AGCTGGCGAG	CGCGGTAGAA	CAGGACTCCC	GTGGTGTAGG	CAGCGGCGAG	AGTCGGGTAA
550	560	570	580	590	600
CCGGCTTTCT	CGGCTGCCGG	CATTTTTCCT	AGGCCAATGC	CTTCCTCGAA	AGCAAAGGCT
GGCCGAAAGA	GCCGACGGCC	GTAAAAAGGG	TCCGGTTACG	GAAGGAGCTT	TCGTTTCCGA

FIG. 8A

610	620	630	640	650	660
TCGATCCGAT	CGACTGGCGG	CTGCCGGAAA	ATCCGGCTGC	GGACATCAAC	TGAAGGCTTG
AGCTAGGCTA	GCTGACCGCC	GACGGCCTTT	TAGGCCGACG	CCTGTAGTTG	ACTTCCGAAC
670	680	690	700	710	720
GCGCGAATGA	CGGCTTTGTC	GTCGCCCTGA	GGTCTTGCCCT	TGGCGGCGGC	GATCCGCCTA
CGCGCTTACT	GCCGAAACAG	CAGCGGGACT	CCAGAACGGA	ACCGCCGCCG	CTAGGCGGAT
730	740	750	760	770	780
AGACGCCC GA	ACGAAATGGC	GGAGGCGGGC	ATGCGCAAAA	TTCTGATCAT	CGGCATCGGT
TCTGCGGGCT	TGCTTTACCG	CCTCCGCCCG	TACGCGTTTT	AAGACTAGTA	GCCGTAGCCA
790	800	810	820	830	840
TCGGGCAATC	CCGAACACAT	GACCGTGCAG	GCGATCAACG	CGCTGAACTG	CGCCGACGTG
AGCCCGTTAG	GGCTTGTGTA	CTGGCACGTC	CGCTAGTTGC	GCGACTTGAC	GCGGCTGCAC
850	860	870	880	890	900
CTCTTTATCC	CGACCAAGGG	AGCGAAGAAG	ACCGAGCTTG	CCGAAGTGCG	CCGCGACATC
GAGAAATAGG	GCTGGTTCCC	TCGCTTCTTC	TGGCTCGAAC	GGCTTCACGC	GGCGCTGTAG
910	920	930	940	950	960
TGCGCCCGCT	ACGTCACGCG	CAAGGACAGC	CGCACCGTCG	AGTTGCGGGT	GCCCGTGCGG
ACGCGGGCGA	TGCAGTGCGC	GTTCTGTGTC	GCGTGGCAGC	TCAAGCGCCA	CGGGCACGCC
970	980	990	1000	1010	1020
CGCACCGAAG	GCGTCAGCTA	TGACGGCAGC	GTCGATGACT	GGCACGCCCA	GATCGCTGGG
GCGTGGCTTC	CGCAGTCGAT	ACTGCCGTCG	CAGCTACTGA	CCGTGCGGGT	CTAGCGACCC
1030	1040	1050	1060	1070	1080
ATTTACGAAG	CGCTTCTATC	GAAGGAGTTG	GGCGAAGAGG	GAAGTGGCGC	GTTTCTCGTC
TAAATGCTTC	GCGAAGATAG	CTTCCTCAAC	CCGCTTCTCC	CTTGACCGCG	CAAAGAGCAG
1090	1100	1110	1120	1130	1140
TGGGGCGACC	CGATGCTCTA	TGACAGCACC	ATTCGCATCG	TCGAGCGGGT	CAAGGCACGC
ACCCCGCTGG	GCTACGAGAT	ACTGTCGTGG	TAAGCGTAGC	AGCTCGCCCA	GTTCCGTGCG
1150	1160	1170	1180	1190	1200
GGTGAGGTCG	CCTTCGCCTA	CGACGTCATT	CCCGGGATCA	CCAGTCTGCA	GGCGCTTTGC
CCACTCCAGC	GGAAGCGGAT	GCTGCAGTAA	GGGCCCTAGT	GGTCAGACGT	CCGCGAAACG

FIG. 8B

1210	1220	1230	1240	1250	1260
GCCAGCCACC	GCATTCCGCT	GAACCTCGTC	GGCAAGCCGG	TGGAGATCAC	CACGGGGCGT
CGGTCGGTGG	CGTAAGGCGA	CTTGGAGCAG	CCGTTCGGCC	ACCTCTAGTG	GTGCCCCGCA
1270	1280	1290	1300	1310	1320
CGGCTGCACG	AAAGCTTTCC	CGAGAAGAGC	CAGACCTCGG	TCGTCATGCT	CGATGGCGAA
GCCGACGTGC	TTTCGAAAGG	GCTCTTCTCG	GTCTGGAGCC	AGCAGTACGA	GCTACCGCTT
1330	1340	1350	1360	1370	1380
CAGGCGTTTC	AGCGGGTCGA	GGACCCGGAG	GCGGAGATCT	ATTGGGGCGC	CTATCTCGGC
GTCCGCAAAG	TCGCCCAGCT	CCTGGGCCTC	CGCCTCTAGA	TAACCCCGCG	GATAGAGCCG
1390	1400	1410	1420	1430	1440
ACGCGGGATG	AGATCGTCAT	TTCCGGCCGC	GTGGCTGAGG	TGAAGGACCG	GATCCTTGAA
TGCGCCCTAC	TCTAGCAGTA	AAGGCCGGCG	CACCGACTCC	ACTTCCTGGC	CTAGGAACCT
1450	1460	1470	1480	1490	1500
ACGCGGGCGG	CGGCGCGCGC	GAAGATGGGA	TGGATCATGG	ACATCTATCT	CCTGCGCAAG
TGCGCCCGCC	GCCGCGCGCG	CTTCTACCCT	ACCTAGTACC	TGTAGATAGA	GGACGCGTTC
1510	1520	1530	1540	1550	1560
GGCGCCGACT	TCGACGAGTG	ACGGGGAGGG	CCGATCTGCG	TCGTGTTTGA	TCTCACTCAA
CCGCGGCTGA	AGCTGCTCAC	TGCCCCTCCC	GGCTAGACGC	AGCACAAACT	AGAGTGAGTT
1570	1580	1590	1600	1610	1620
GGTTTGCGGC	TGTGTTATAG	CGTCTTAAGA	GGCTTCTTCA	GGGAGGAGAA	CCTCAAGTGA
CCAAACGCCG	ACACAATATC	GCAGAAATTCT	CCGAAGAAGT	CCCTCCTCTT	GGAGTTCACT
1630	1640	1650	1660	1670	1680
TGACGGATTT	GATGACCAGC	TGCGCCCTTC	CATTGACCGG	AGATGCCGGC	ACCGTCGCTT
ACTGCCTAAA	CTACTGGTCG	ACGCGGGAAG	GTAAGTGGCC	TCTACGGCCG	TGGCAGCGAA
1690	1700	1710	1720	1730	1740
CGATGCGCCG	CGGCGCCTGC	CCGTCCTTGG	CAGAGCCGAT	GCAGACCGGC	GACGGCCTGC
GCTACGCGGC	GCCGCGGACG	GGCAGGAACC	GTCTCGGCTA	CGTCTGGCCG	CTGCCGGACG
1750	1760	1770	1780	1790	1800
TCGTGAGGGT	GAGGCCAACG	GATGACAGCC	TGACGCTGCC	GAAGGTCATT	GCCCTTGCCA
AGCACTCCCA	CTCCGGTTGC	CTACTGTCTG	ACTGCGACGG	CTTCCAGTAA	CGGGAACGGT

FIG. 8C

1810	1820	1830	1840	1850	1860
CGGCTGCCGA	GCGCTTCGGC	AATGGCATCA	TCGAGATTAC	CGCGCGCGGA	AACCTGCAGC
GCCGACGGCT	CGCGAAGCCG	TTACCGTAGT	AGCTCTAATG	GCGCGCGCCT	TTGGACGTGC
1870	1880	1890	1900	1910	1920
TTCGCGGCCT	GAGCGCGGCT	TCGGTGCCAA	GGCTGGCGCA	GGCGATCGGC	GATGCGGAGA
AAGCGCCGGA	CTCGCGCCGA	AGCCACGGTT	CCGACCGCGT	CCGCTAGCCG	CTACGCCTCT
1930	1940	1950	1960	1970	1980
TCGCCATTGC	CGAGGGGCTC	GCGATCGAGG	TGCCGCCCCT	GGCCGGCATC	GACCCGGACG
AGCGGTAACG	GCTCCCCGAG	CGCTAGCTCC	ACGGCGGGGA	CCGGCCGTAG	CTGGGCCTGC
1990	2000	2010	2020	2030	2040
AGATCGCCGA	TCCGCGGCCG	ATTGCCACTG	AGCTTCGTGA	AGCGTTGGAT	GTGCGCCAGG
TCTAGCGGCT	AGGCGCCGGC	TAACGGTGAC	TCGAAGCACT	TCGCAACCTA	CACGCGGTCC
2050	2060	2070	2080	2090	2100
TGCCGTTGAA	GCTTGACACC	AAATTATCCG	TCGTCATCGA	TAGCGGTGGC	CGGTTTGGTC
ACGGCAACTT	CGAACGTGGG	TTTAATAGGC	AGCAGTAGCT	ATCGCCACCG	GCCAAACCAG
2110	2120	2130	2140	2150	2160
TCGGCGCTGT	CGTCGCCGAC	ATTGCGCTTC	AGGCGGTTTC	GACTGTCGCG	GGGGTGGCCT
AGCCGCGACA	GCAGCGGCTG	TAAGCGGAAG	TCCGCCAAAG	CTGACAGCGC	CCCCACCGGA
2170	2180	2190	2200	2210	2220
GGGTGCTGTC	GCTTGGCGGC	ACGTCAACGA	AGGCATCGAG	CGTCGGGACG	TTGGCCGGCA
CCCACGACAG	CGAACCGCCG	TGCAGTTGCT	TCCGTAGCTC	GCAGCCCTGC	AACCGGCCGT
2230	2240	2250	2260	2270	2280
ACGCGGTCGT	GCCGGCCCTG	ATCACCATTG	TCGAGAAACT	GGCGAGCCTG	GGCACGACGA
TGCGCCAGCA	CGGCCGGGAC	TAGTGGTAAG	AGCTCTTTGA	CCGCTCGGAC	CCGTGCTGCT
2290	2300	2310	2320	2330	2340
TGCGCGGGCG	CGATCTGGAC	CCGTGCGGAA	TCCGCGCGCT	CTGTGCTGCT	GAGACATCGT
ACGCGCCCCG	GCTAGACCTG	GGCAGCCTTT	AGGCGCGCGA	GACAGCGACA	CTCTGTAGCA
2350	2360	2370	2380	2390	2400
CCGAACGCC	GGCCGCTCCG	CGTTCGGCCG	CAATACCCGG	CATTCATGCG	CTGGGTAACG
GGCTTGCGGG	CCGGCGAGGC	GCAAGCCGGC	GTTATGGGCC	GTAAGTACGC	GACCCATTGC

FIG. 8D

2410	2420	2430	2440	2450	2460
CCGACACCGT	TCTCGGCCTC	GGTCTGGCCT	TTGCTCAGGT	GGAGGCCGCC	GCGCTGGCAT
GGCTGTGGCA	AGAGCCGGAG	CCAGACCGGA	AACGAGTCCA	CCTCCGGCGG	CGCGACCGTA
2470	2480	2490	2500	2510	2520
CCTACCTGCA	TCAGGTCCAG	GCGCTTGGCG	CCAATGCGAT	CCGGCTTGCG	CCCGGGCACG
GGATGGACGT	AGTCCAGGTC	CGCGAACCGC	GGTTACGCTA	GGCCGAACGC	GGGCCCCTGC
2530	2540	2550	2560	2570	2580
CCTTCTTCGT	CCTCGGCCTT	TGCCCCGAGA	CCGCGGCTGT	GGCGCAGAGC	CTGGCAGCGT
GGAAGAAGCA	GGAGCCGGAA	ACGGGGCTCT	GGCGCCGACA	CCGCGTCTCG	GACCGTCGCA
2590	2600	2610	2620	2630	2640
CACACGTTTT	TCGCATTGCC	GAGCAGGATC	CGCGCAATGC	GATCGCCACC	TGCGCCGGCA
GTGTGCCAAA	AGCGTAACGG	CTCGTCCTAG	GCGCGTTACG	CTAGCGGTGG	ACGCGGCCGT
2650	2660	2670	2680	2690	2700
GCAAGGGTTG	CGCCTCGGCG	TGGATGGAAA	CCAAGGGCAT	GGCCGAGCGC	CTCGTCGAGA
CGTTCCCAAC	GCGGAGCCGC	ACCTACCTTT	GGTTCCCGTA	CCGGCTCGCG	GAGCAGCTCT
2710	2720	2730	2740	2750	2760
CGGCGCCGGA	ATTGCTCGAC	GGGTCGCTCA	CCGTGCATCT	CTCCGGCTGC	GCCAAGGGCT
GCCGCGGCCT	TAACGAGCTG	CCCAGCGAGT	GGCACGTAGA	GAGGCCGACG	CGGTTCCCGA
2770	2780	2790	2800	2810	2820
GCGCCCGGCC	GAAGCCGTCC	GAAGTACGCG	TTGTGCGTGC	GCCATCAGGA	TACGGGCTTG
CGCGGGCCGG	CTTCGGCAGG	CTTGACTGCG	AACAGCCACG	CGGTAGTCCT	ATGCCCCAAC
2830	2840	2850	2860	2870	2880
TCGTAAATGG	GGCTGCCAAT	GGCTTGCCAA	GCGCCTACAC	CGATGAGAAT	GGAATGGGAT
AGCATTTACC	CCGACGGTTA	CCGAACGGTT	CGCGGATGTG	GCTACTCTTA	CCTTACCCTA
2890	2900	2910	2920	2930	2940
CCGCCCTTGC	CCGGCTCGGC	CGGCTGGTGC	GGCAAAACAA	AGACGCTGGC	GAATCGGCGC
GGCGGGAACG	GGCCGAGCCG	GCCGACCACG	CCGTTTTGTT	TCTGCGACCG	CTTAGCCGCG
2950	2960	2970	2980	2990	3000
AGTCCTGTCT	TACACGGCTC	GGAGCTGCGC	GCGTCTCGGC	AGCGTTCGAA	CAGGGATAGA
TCAGGACAGA	ATGTGCCGAG	CCTCGACGCG	CGCAGAGCCG	TCGCAAGCTT	GTCCCTATCT

FIG. 8E

3010	3020	3030	3040	3050	3060
CATGCCTGAG	TATGATTACA	TTCGCGATGG	CAACGCCATC	TACGAGCGTT	CCTTCGCCAT
GTACGGACTC	ATACTAATGT	AAGCGCTACC	GTTGCGGTAG	ATGCTCGCAA	GGAAGCGGTA
3070	3080	3090	3100	3110	3120
CATCCGCGCC	GAGGCCGATC	TGTCGCGCTT	CTCCGAAGAG	GAAGCGGATC	TGGCTGTGCG
GTAGGCGCGG	CTCCGGCTAG	ACAGCGCGAA	GAGGCTTCTC	CTTCGCCTAG	ACCGACACGC
3130	3140	3150	3160	3170	3180
CATGGTGCAC	GCCTGCGGTT	CCGTGCGAGG	GACCAGGCAG	TTCGTGTTTT	CTCCCGATTT
GTACCACGTG	CGGACGCCAA	GGCAGCTCCG	CTGGTCCGTC	AAGCACAAAA	GAGGGCTAAA
3190	3200	3210	3220	3230	3240
CGTAAGCTCG	GCCCCGTGCGG	CGCTGAAAGC	CGGTGCGCCG	ATCCTCTGCG	ATGCCGAGAT
GCATTCGAGC	CGGGCACGCC	GCGACTTTTC	GCCACGCGGC	TAGGAGACGC	TACGGCTCTA
3250	3260	3270	3280	3290	3300
GGTTGCGCAC	GGTGTCACCC	GCGCCCGTCT	GCCGGCCGGC	AACGAGGTGA	TCTGCACGCT
CCAACGCGTG	CCACAGTGGG	CGCGGGCAGA	CGGCCGGCCG	TTGCTCCACT	AGACGTGCGA
3310	3320	3330	3340	3350	3360
GCGCGATCCT	CGCACGCCCC	CACTTGCGGC	CGAGATCGGC	AACACCCGCT	CCGCCGCAGC
CGCGCTAGGA	GCGTGCGGGC	GTGAACGCCG	GCTCTAGCCG	TTGTGGGCGA	GGCGGCGTCG
3370	3380	3390	3400	3410	3420
CCTGAAGCTC	TGGAGCGAGC	GGCTGGCCGG	TTCGGTGGTC	GCGATCGGCA	ACGCGCCGAC
GGACTTCGAG	ACCTCGCTCG	CCGACCGGCC	AAGCCACCAG	CGCTAGCCGT	TGCGCGGCTG
3430	3440	3450	3460	3470	3480
GGCGTTGTTC	TTCCTCTTGG	AAATGCTGCG	CGACGGCGCG	CCGAAGCCGG	CGGCAATCCT
CCGCAACAAG	AAGGAGAACC	TTTACGACGC	GCTGCCGCGC	GGCTTCGGCC	GCCGTTAGGA
3490	3500	3510	3520	3530	3540
CGGCATGCCC	GTCGGTTTCG	TCGGTGCGGC	GGAATCGAAG	GATGCGCTGG	CCGAGAACTC
GCCGTACGGG	CAGCCAAAGC	AGCCACGCCG	CCTTAGCTTC	CTACGCGACC	GGCTCTTGAG
3550	3560	3570	3580	3590	3600
CTATGGCGTT	CCCTTCGCCA	TCGTGCGCGG	CCGCCTCGGC	GGGAGTGCCA	TGACGGCGGC
GATACCGCAA	GGGAAGCGGT	AGCACGCGCC	GGCGGAGCCG	CCCTCACGGT	ACTGCCGCCG

FIG. 8F

3610	3620	3630	3640	3650	3660
AGCGCTTAAC	TCGCTCGCGA	GGCCGGGCCT	GTGAGCGGCG	TCGGCGTGGG	GCGCCTGATC
TCGCGAATTG	AGCGAGCGCT	CCGGCCCCGA	CACTCGCCGC	AGCCGCACCC	CGCGGACTAG
3670	3680	3690	3700	3710	3720
GGTGTGGA	CCGGCCCCGG	TGATCCGGAA	CTTTTGACGG	TCAAGGCGGT	GAAGGCGCTC
CCACAACCCT	GGCCGGGGCC	ACTAGGCCTT	GAAAACTGCC	AGTTCCGCCA	CTTCCGCGAG
3730	3740	3750	3760	3770	3780
GGGCAAGCCG	ATGTGCTTGC	CTATTTTCGCC	AAGGCCGGGC	GAAGCGGTAA	CGGCCGCGCG
CCCGTTCCGG	TACACGAACG	GATAAAGCGG	TTCCGGCCCC	CTTCGCCATT	GCCGGCAGCG
3790	3800	3810	3820	3830	3840
GTGGTCGAGG	GTCTGCTGAA	GCCCGATCTT	GTCGAGCTGC	CGCTATACTA	TCCGGTGACG
CACCAGCTCC	CAGACGACTT	CGGGCTAGAA	CAGCTCGACG	GCGATATGAT	AGGCCACTGC
3850	3860	3870	3880	3890	3900
ACCGAAATCG	ACAAGGACGA	TGGCGCCTAC	AAGACCCAGA	TCACCGACTT	CTACAATGCG
TGGCTTTAGC	TGTTCCCTGCT	ACCGCGGATG	TTCTGGGTCT	AGTGGCTGAA	GATGTTACGC
3910	3920	3930	3940	3950	3960
TCGGCCGAAG	CGGTAGCGGC	GCATCTTGCC	GCCGGGCGCA	CGGTCGCCGT	GCTCAGTGAA
AGCCGGCTTC	GCCATCGCCG	CGTAGAACGG	CGGCCCGCGT	GCCAGCGGCA	CGAGTCACTT
3970	3980	3990	4000	4010	4020
GGCGACCCGC	TGTTCTATGG	TTCCTACATG	CATCTGCATG	TGCGGCTCGC	CAATCGTTTC
CCGCTGGGCG	ACAAGATACC	AAGGATGTAC	GTAGACGTAC	ACGCCGAGCG	GTTAGCAAAG
4030	4040	4050	4060	4070	4080
CCGGTCGAGG	TGATCCCCGG	CATTACCGCC	ATGTCCGGCT	GTTGGTCGCT	TGCCGGCCTG
GGCCAGCTCC	ACTAGGGGCC	GTAATGGCGG	TACAGGCCGA	CAACCAGCGA	ACGGCCGGAC
4090	4100	4110	4120	4130	4140
CCGCTGGTGC	AGGGCGACGA	CGTGCTCTCG	GTGCTTCCGG	GCACCATGGC	CGAGGCCGAG
GGCGACCACG	TCCCCTGCTG	GCACGAGAGC	CACGAAGGCC	CGTGGTACCG	GCTCCGGCTC
4150	4160	4170	4180	4190	4200
CTCGGCCGCA	GGCTTGCGGA	TACCGAAGCC	GCCGTGATCA	TGAAGGTCGG	GCGCAATTTG
GAGCCGGCGT	CCGAACGCCT	ATGGCTTCGG	CGGCACTAGT	ACTTCCAGCC	CGCGTTAAAC

FIG. 8G

4210	4220	4230	4240	4250	4260
CCGAAGATCC	GTCGGGCGCT	CGCTGCCTCC	GGCCGTCTCG	ACCAGGCCGT	CTATGTCGAA
GGCTTCTAGG	CAGCCCGCGA	GCGACGGAGG	CCGGCAGAGC	TGGTCCGGCA	GATACAGCTT
4270	4280	4290	4300	4310	4320
CGCGGCACGA	TGAAGAACGC	GGCGATGACG	GCTCTTGCGG	AAAAGGCCGA	CGACGAGGCG
GCGCCGTGCT	ACTTCTTGCG	CCGCTACTGC	CGAGAACGCC	TTTTCCGGCT	GCTGCTCCGC
4330	4340	4350	4360	4370	4380
CCCTATTTCT	CGCTGGTGCT	CGTTCCCGGC	TGGAAGGACC	GACCATGACC	GGTACGCTCT
GGGATAAAGA	GCGACCACGA	GCAAGGGCCG	ACCTTCCTGG	CTGGTACTGG	CCATGCGAGA
4390	4400	4410	4420	4430	4440
ATGTCGTCGG	TACCGGACCG	GGCAGCGCCA	AGCAGATGAC	GCCGGAAACG	GCGGAAGCCG
TACAGCAGCC	ATGGCCTGGC	CCGTCGCGGT	TCGTCTACTG	CGGCCTTTGC	CGCCTTCGGC
4450	4460	4470	4480	4490	4500
TTGCGGCCGC	TCAGGAGTTT	TACGGCTACT	TTCCCTATCT	CGACCGGCTG	AACCTCAGAC
AACGCCGGCG	AGTCCTCAAA	ATGCCGATGA	AAGGGATAGA	GCTGGCCGAC	TTGGAGTCTG
4510	4520	4530	4540	4550	4560
CGGATCAGAT	CCGTGTCGCC	TCGGACAACC	GCGAGGAGCT	CGATCGGGCA	CAGGTCGCGC
GCCTAGTCTA	GGCACAGCGG	AGCCTGTTGG	CGCTCCTCGA	GCTAGCCCGT	GTCCAGCGCG
4570	4580	4590	4600	4610	4620
TGACGCGGGC	TGCGGCAGGC	GTGAAGGTCT	GCATGGTCTC	CGGTGGCGAT	CCCGGTGTCT
ACTGCGCCCG	ACGCCGTCCG	CACTTCCAGA	CGTACCAGAG	GCCACCGCTA	GGGCCACAGA
4630	4640	4650	4660	4670	4680
TTGCCATGGC	GGCCGCCGTC	TGCGAGGCGA	TCGACAAGGG	ACCGGCGGAA	TGGAAGTCGG
AACGGTACCG	CCGGCAGCAG	ACGCTCCGCT	AGCTGTTCCC	TGGCCGCCTT	ACCTTCAGCC
4690	4700	4710	4720	4730	4740
TTGAACTGGT	GATCACGCCC	GGCGTGACCG	CGATGCTCGC	CGTTGCCGCC	CGCATCGGCG
AACTTGACCA	CTAGTGCGGG	CCGCACTGGC	GCTACGAGCG	GCAACGGCGG	GCGTAGCCGC
4750	4760	4770	4780	4790	4800
CGCCGCTCGG	TCATGATTTT	TGTGCGATCT	CGCTTTCCGA	CAATCTGAAG	CCCTGGGAAG
GCGGCGAGCC	AGTACTAAAG	ACACGCTAGA	GCGAAAGGCT	GTTAGACTTC	GGGACCCTTC

FIG. 8H

4810	4820	4830	4840	4850	4860
TCATCACCCG	GCGTCTCAGG	CTGGCGGCGG	AAGCGGGCTT	CGTCATTGCC	CTCTACAATC
AGTAGTGGGC	CGCAGAGTCC	GACCGCCGCC	TTCGCCCCGAA	GCAGTAACGG	GAGATGTTAG
4870	4880	4890	4900	4910	4920
CGATCAGCAA	GGCGCGGCC	TGGCAGCTCG	GTGAGGCCTT	CGAGCTTCTG	CGCAGCGTTC
GCTAGTCGTT	CCGCGCCGGG	ACCGTCGAGC	CACTCCGGAA	GCTCGAAGAC	GCGTCGCAAG
4930	4940	4950	4960	4970	4980
TGCCGGCAAG	CGTTCCGGTC	ATCTTCGGCC	GTGCGGCCGG	GCGGCCGGAC	GAACGGATCG
ACGGCCGTTC	GCAAGGCCAG	TAGAAGCCGG	CACGCCGGCC	CGCCGGCCTG	CTTGCCCTAGC
4990	5000	5010	5020	5030	5040
CGGTGATGCC	GCTCGGCGAG	GCCGATGCCA	ACCGCGCCGA	CATGGCGACC	TGCGTCATCA
GCCACTACGG	CGAGCCGCTC	CGGCTACGGT	TGGCGCGGCT	GTACCGCTGG	ACGCAGTAGT
5050	5060	5070	5080	5090	5100
TCGGCTCGCC	GGAGACGCGC	ATCGTCGAGC	GCGACGGCCA	ACCCGATCTC	GTCTACACAC
AGCCGAGCGG	CCTCTGCGCG	TAGCAGCTCG	CGCTGCCGGT	TGGGCTAGAG	CAGATGTGTG
5110	5120	5130	5140	5150	5160
CGCGCTTCTA	TGCAGGGGCG	AGCCAGTGAG	CGATGCGGTT	GAGTGCCTCG	TCGCAACTGC
GCGCGAAGAT	ACGTCCCCGC	TCGGTCACTC	GCTACGCCAA	CTCACGGAGC	AGCGTTGACG
5170	5180	5190	5200	5210	5220
CGACCGTCGG	CACGTCCGCG	GGCTTGCGCC	GCTCGACCAT	GATCACCTCG	ATGCCGAGCC
GCTGGCAGCC	GTGCAGGCGC	CCGAACGCGG	CGAGCTGGTA	CTAGTGAGAG	TACGGCTCGG
5230	5240	5250	5260	5270	5280
GGCGCGCTGC	GGCAATCTTG	CCGTAGGTGG	CGCTGCCACC	GCTGTTCTTG	GCGACGATCA
CCGCGCGACG	CCGTTAGAAC	GGCATCCACC	GCGACGGTGG	CGACAAGAAC	CGCTGCTAGT
5290	5300	5310	5320	5330	5340
CATCGATCTG	CCGACTCCTG	AGCAACGCGG	CTTCGTCGGC	TTCCGCAAAG	GGACCGGTCTG
GTAGCTAGAC	GGCTGAGGAC	TCGTTGCGCC	GAAGCAGCCG	AAGGCGTTTC	CCTGGCCAGC
5350	5360	5370	5380	5390	5400
CCAGGATCGC	CTCCTGGTCG	GGCAGATTAA	GCGGCGGCGT	CACCGGATCG	ACGCTGCGGA
GGTCCTAGCG	GAGGACCAGC	CCGTCTAATT	CGCCGCCGCA	GTGGCCTAGC	TGCGACGCCT

FIG. 8I

5410	5420	5430	5440	5450	5460
TGACGTAGCT	GTGCTGCGGC	GCGACCTCGA	AGTGGAAAGC	TTCCTGTCTGA	CCTATCGCCA
ACTGCATCGA	CACGACGCCG	CGCTGGAGCT	TCACCTTTTCG	AAGGACAGCT	GGATAGCGGT
5470	5480	5490	5500	5510	5520
GGAAGACGCG	GCGTCGCCGA	TCACCGAGCG	CGCTGACGGC	CTCGACAACG	CTATCGACAG
CCTTCTGCGC	CGCAGCGGCT	AGTGGCTCGC	GCGACTGCCG	GAGCTGTTGC	GATAGCTGTC
5530	5540	5550	5560	5570	5580
CAGTCCAGCG	GTCGCCAGGC	AGGGGCACCC	ATTCCGGTCG	GCGGAGGGCG	ATAAGCGCAA
GTCAGGTGCG	CAGCGGTCCG	TCCCCGTGGG	TAAGGCCAGC	CGCCTCCCGC	TATTCGCGTT
5590	5600	5610	5620	5630	5640
CGCCGGTTCT	TTGCGCTGCG	TCCGCGGCGT	TGTGCGAAAT	GCGTGCGGCA	AAGGGGTGCG
GCGGCCAAGA	AACGCGACGC	AGGCGCCGCA	ACACGCTTTA	CGCACGCCGT	TTCCCCACGC
5650	5660	5670	5680	5690	5700
TCGCATCGAC	CAGCAGCGCG	ATGTTTTCTGT	CATGCACGAA	ATGCGCCAGC	CCATCGGCGC
AGCGTAGCTG	GTCGTCGCGC	TACAAAAGCA	GTACGTGCTT	TACGCGGTCTG	GGTAGCCGCG
5710	5720	5730	5740	5750	5760
CGCCAAAGCC	GCCGATGCGC	GTCTTGACCG	GCTGCGGCCG	CGGGTCCGCG	GTGCGGCCGG
GCGGTTTCGG	CGGCTACGCG	CAGAACTGGC	CGACGCCGGC	GCCCAGGCGC	CACGCCGGCC
5770	5780	5790	5800	5810	5820
CCAGCGAGAT	GGCGGTGTCTG	TAGCGGACAT	CTTCGGCCAA	GCGGCGCGCG	AGTTCGCGTG
GGTCGCTCTA	CCGCCACAGC	ATCGCCTGTA	GAAGCCGGTT	CGCCGCGCGC	TCAAGCGCAC
5830	5840	5850	5860	5870	5880
CCTCGGTGGT	GCCACCCAGA	ATCAGAATAC	GAGGTTTTTC	CATGGCTGAC	GTGTCGAACA
GGAGCCACCA	CGGTGGGTCT	TAGTCTTATG	CTCCAAAAAG	GTACCGACTG	CACAGCTTGT
5890	5900	5910	5920	5930	5940
GCGAACCCGC	CATAGTCTCC	CCCTGGCTGA	CCGTCATCGG	TATCGGTGAG	GATGGTGTAG
CGCTTGGGCG	GTATCAGAGG	GGGACCGACT	GGCAGTAGCC	ATAGCCACTC	CTACCACATC
5950	5960	5970	5980	5990	6000
CGGGTCTCGG	CGACGAGGCC	AAGCGGCTGA	TCGCCGAAGC	GCCGGTCGTC	TACGGCGGCC
GCCCAGAGCC	GCTGCTCCGG	TTCGCCGACT	AGCGGCTTCG	CGGCCAGCAG	ATGCCGCCGG

FIG. 8J

6010	6020	6030	6040	6050	6060
ATCGTCATCT	GGAGCTCGCC	GCCTCCCTCA	TCACCGGCCGA	AGCGCACAAT	TGGCTAAGCC
TAGCAGTAGA	CCTCGAGCGG	CGGAGGGAGT	AGTGGCCGCT	TCGCGTGTTA	ACCGATTTCGG
6070	6080	6090	6100	6110	6120
CCCTCGAACG	CTCGGTCGTC	GAGATCGTCG	CGCGTCGCGG	CAGCCCGGTG	GTGGTGCTTG
GGGAGCTTGC	GAGCCAGCAG	CTCTAGCAGC	GCGCAGCGCC	GTCGGGCCAC	CACCACGAAC
6130	6140	6150	6160	6170	6180
CCTCGGGCGA	CCCGTTCTTC	TTCGGCGTCG	GCGTGACGCT	GGCGCGCCGC	ATCGCCTCGG
GGAGCCCGCT	GGGCAAGAAG	AAGCCGCAGC	CGCACTGCGA	CCGCGCGGCG	TAGCGGAGCC
6190	6200	6210	6220	6230	6240
CCGAAATACG	CACGCTTCCG	GCGCCGTCGT	CGATCAGTCT	TGCCGCCTCG	CGCCTCGGCT
GGCTTTATGC	GTGCGAAGGC	CGCGGCAGCA	GCTAGTCAGA	ACGGCGGAGC	GCGGAGCCGA
6250	6260	6270	6280	6290	6300
GGGCGCTGCA	GGATGCGACG	CTCGTCTCCG	TACATGGGCG	GCCGCTGGAT	CTGGTGCGAC
CCC GCGACGT	CCTACGCTGC	GAGCAGAGGC	ATGTACCCGC	CGGCGACCTA	GACCACGCTG
6310	6320	6330	6340	6350	6360
CGCATTTGCA	TCCGGGGGCG	CGTGTGCTTA	CGCTCACGTC	GGACGGTGCG	GGTCCGCGAG
GCGTAAACGT	AGGCCCCCGC	GCACACGAAT	GCGAGTG CAG	CCTGCCACGC	CCAGGCGCTC
6370	6380	6390	6400	6410	6420
ACCTTGCCGA	GCTTCTGGTT	TCAAGCGGCT	TCGGTCAGTC	GCGACTGACC	GTGCTCGAAG
TGGAACGGCT	CGAAGACCAA	AGTTCGCCGA	AGCCAGTCAG	CGCTGACTGG	CACGAGCTTC
6430	6440	6450	6460	6470	6480
CGCTGGGCGG	CGCCGGCGAA	CGGGTGACGA	CGCAGATCGC	CGCGCGCTTC	ATGCTCGGCC
GCGACCCGCC	GCGGCCGCTT	GCCCACTGCT	GCGTCTAGCG	GCGCGCGAAG	TACGAGCCGG
6490	6500	6510	6520	6530	6540
TCGTGCATCC	TTTGAACGTC	TGCGCCATTG	AGGTGGCGGC	CGACGAGGGC	GCGCGCATCC
AGCACGTAGG	AAACTTGCAG	ACGCGGTAAC	TCCACCGCCG	GCTGCTCCCG	CGCGCGTAGG
6550	6560	6570	6580	6590	6600
TGCCGCTTGC	CGCCGGCCGC	GACGATGCGC	TGTTCAACA	TGACGGGCAG	ATCACCAAGC
ACGGCGAACG	GCGGCCGGCG	CTGCTACGCG	ACAAGCTTGT	ACTGCCCGTC	TAGTGTTTCG

FIG. 8K

6610	6620	6630	6640	6650	6660
GCGAGGTGCG	GGCGCTGACG	CTGTCCGGAC	TCGCACCGCG	CAAGGGCGAA	CTGCTATGGG
CGCTCCACGC	CCGCGACTGC	GACAGCCGTG	AGCGTGGCGC	GTTCCCGCTT	GACGATACCC
6670	6680	6690	6700	6710	6720
ACATCGGCGG	CGGCTCCGGC	TCGATCGGCA	TCGAATGGAT	GCTCGCCGAT	CCGACCATGC
TGTAGCCGCC	GCCGAGGCCG	AGCTAGCCGT	AGCTTACCTA	CGAGCGGCTA	GGCTGGTACG
6730	6740	6750	6760	6770	6780
AGGCGATCAC	CATCGAGGTT	GAGCCGGAGC	GGGCAGCGCG	CATCGGCCGC	AACGCGACGA
TCCGCTAGTG	GTAGCTCCAA	CTCGGCCTCG	CCCGTCGCGC	GTAGCCGGCG	TTGCGCTGCT
6790	6800	6810	6820	6830	6840
TGTTTCGGCGT	GCCCCGGGCTG	ACGGTTGTCTG	AAGGCGAGGC	GCCGGCGGCG	CTTGCCGGCC
ACAAGCCGCA	CGGGCCCGAC	TGCCAACAGC	TTCCGCTCCG	CGGCCGCCGC	GAACGGCCGG
6850	6860	6870	6880	6890	6900
TGCCACAACC	GGACGCGATC	TTCATCGGCG	GCGGCGGCAG	CGAAGACGGC	GTCATGGAAG
ACGGTGTGTTG	CCTGCGCTAG	AAGTAGCCGC	CGCCGCCGTC	GCTTCTGCCG	CAGTACCTTC
6910	6920	6930	6940	6950	6960
CAGCGATCGA	GGCGCTCAAG	TCAGGCGGAC	GGCTGGTTGC	CAACGCGGTG	ACGACGGACA
GTCGCTAGCT	CCGCGAGTTC	AGTCCGCCTG	CCGACCAACG	GTTGCGCCAC	TGCTGCCTGT
6970	6980	6990	7000	7010	7020
TGGAAGCGGT	GCTGCTCGAT	CATCACGCGC	GGCTCGGCGG	TTCGCTGATC	CGCATCGATA
ACCTTCGCCA	CGACGAGCTA	GTAGTGCGCG	CCGAGCCGCC	AAGCGACTAG	GCGTAGCTAT
7030	7040	7050	7060	7070	7080
TCGCGCGTGC	AGGACCCATC	GGCGGCATGA	CCGGCTGGAA	GCCGGCCATG	CCGGTCACCC
AGCGCGCACG	TCCTGGGTAG	CCGCCGTACT	GGCCGACCTT	CGGCCGGTAC	GGCCAGTGGG
7090	7100	7110	7120	7130	7140
AATGGTCGTG	GACGAAGGGC	TAAAGCAGTT	CCAGCGAAAG	TGTGACGCGG	TTTTGCGTCC
TTACCAGCAC	CTGCTTCCCG	ATTTCTGTC	GGTCGCTTTC	ACACTGCGCC	AAAACGCAGG
7150	7160	7170	7180	7190	7200
GGAAGTGC	AAGAAAAAGA	AAGAGTAACC	TATGACGGTA	CATTTTCATCG	GCGCCGGCCC
CCTTGACGCG	TTCTTTTTCT	TTCTCATTGG	ATACTGCCAT	GTAAAGTAGC	CGCGGCCGGG

FIG. 8L

7210	7220	7230	7240	7250	7260
GGGAGCCGCA	GACCTGATCA	CGGTGCGTGG	TCGCGACCTG	ATCGGGCGCT	GCCCCGGTCTG
CCCTCGGCGT	CTGGACTAGT	GCCACGCACC	AGCGCTGGAC	TAGCCCCGCGA	CGGGCCAGAC
7270	7280	7290	7300	7310	7320
CCTTTACGCC	GGCTCGATCG	TCTCGCCGGA	GCTGCTGCGA	TATTGCCCCG	CGGGCGCCCC
GGAAATGCGG	CCGAGCTAGC	AGAGCGGCCT	CGACGACGCT	ATAACGGGCG	GCCCCGCGGC
7330	7340	7350	7360	7370	7380
CATTGTCGAT	ACGGCGCCGA	TGTCCCTCGA	CGAGATCGAG	GCGGAGTATG	TGAAGGCCGA
GTAACAGCTA	TGCCGCGGCT	ACAGGGAGCT	GCTCTAGCTC	CGCCTCATAC	ACTTCCGGCT
7390	7400	7410	7420	7430	7440
AGCCGAAGGG	CTCGACGTGG	CGCGGCTTCA	TTCGGGCGAC	CTTTCGGTCT	GGAGTGCTGT
TCGGCTTCCC	GAGCTGCACC	GCGCCGAAGT	AAGCCCGCTG	GAAAGCCAGA	CCTCACGACA
7450	7460	7470	7480	7490	7500
GGCCGAACAG	ATCCGCCGGC	TCGAGAAGCA	TGGCATCGCC	TATACGATGA	CGCCGGGCGT
CCGGCTTGTC	TAGGCGGCCG	AGCTCTTCGT	ACCGTAGCGG	ATATGCTACT	GCGGCCCGCA
7510	7520	7530	7540	7550	7560
TCCTTCCTTT	GCGGCGGCGG	CTTCAGCGCT	CGGTGCGGAA	TTGACCATTG	CGGCCGTGGC
AGGAAGGAAA	CGCCGCCGCC	GAAGTCGCGA	GCCAGCGCTT	AACTGGTAAG	GCCGGCACCG
7570	7580	7590	7600	7610	7620
CCAGAGCCTG	GTGCTGACCC	GCGTTTCGGG	CCGCGCCTCG	CCGATGCCGA	ACTCAGAAAC
GGTCTCGGAC	CACGACTGGG	CGCAAAGCCC	GGCGCGGAGC	GGCTACGGCT	TGAGTCTTTG
7630	7640	7650	7660	7670	7680
GCTTTCCGCT	TTCGGCGCTA	CGGGATCGAC	GCTGGCAATC	CACCTTGCGA	TCCATGCGCT
CGAAAGGCGA	AAGCCGCGAT	GCCCTAGCTG	CGACCGTTAG	GTGGAACGCT	AGGTACGCGA
7690	7700	7710	7720	7730	7740
TCAGCAGGTG	GTCGAGGAAC	TGACGCCGCT	CTACGGTGCC	GACTGCCCGG	TCGCCATCGT
AGTCGTCCAC	CAGCTCCTTG	ACTGCGGCGA	GATGCCACGG	CTGACGGGCC	AGCGGTAGCA
7750	7760	7770	7780	7790	7800
CGTCAAGGCC	TCCTGGCCGG	ACGAACGCGT	GGTGCGCGGC	ACGCTCGGTG	ACATCGCCCG
GCAGTTCCGG	AGGACCGGCC	TGCTTGCGCA	CCACGCGCCG	TGCGAGCCAC	TGTAGCGGCG

FIG. 8M

7810	7820	7830	7840	7850	7860
CAAGGTGGCG	GAAGAGCCGA	TCGAGCGCAC	GGCGCTGATC	TTCGTCGGTC	CGGGGCTCGA
GTTCCACCGC	CTTCTCGGCT	AGCTCGCGTG	CCGCGACTAG	AAGCAGCCAG	GCCCCGAGCT
7870	7880	7890	7900	7910	7920
AGCCTCCGAT	TTCCGTGAAA	GCTCGCTCTA	CGATCCCGCC	TATCAGCGGC	GCTTCAGAGG
TCGGAGGCTA	AAGGCACTTT	CGAGCGAGAT	GCTAGGGCGG	ATAGTCGCCG	CGAAGTCTCC
7930	7940	7950	7960	7970	7980
GCGCGGCGAA	TAGGCCGCAC	TCCCTCGGGG	GTCGGCCTAA	GTTTCCCGCT	GAGAGGGTTT
CGCGCCGCTT	ATCCGGCGTG	AGGGAGCCCC	CAGCCGGATT	CAAAGGGCGA	CTCTCCCAA
7990	8000	8010	8020	8030	8040
TGAAACCTAT	TCTGCCGGTT	CTTCGCGCGG	CGGCCGCTGC	TTGAGCGGGA	CGCCGCGCTT
ACTTTGGATA	AGACGGCCAA	GAAGCGCGCC	GCCGGCGACG	AACTCGCCCT	GCGGCGCGAA
8050	8060	8070	8080	8090	8100
TTCCTCGACG	CGGTCGCGGT	AGAGCGCTGC	CTGTCCAAGC	AGCATCAGCG	TCACCGGCGT
AAGGAGCTGC	GCCAGCGCCA	TCTCGCGACG	GACAGGTTTC	TCGTAGTCGC	AGTGGCCGCA
8110	8120	8130	8140	8150	8160
GGTGGCGACG	ACGAAGACGA	TGATCAGGAT	TTCGTGGAAT	ACCCAGCGGC	TCTGCAGCAC
CCACCGCTGC	TGCTTCTGCT	ACTAGTCCTA	AAGCACCTTA	TGGGTCGCCG	AGACGTCGTG
8170	8180	8190	8200	8210	8220
GGCAAAGCAG	ATGATAGAGG	CGGCGCAGAT	CATCAGTACG	CCGCCGCTGG	TCGCCAGCGT
CCGTTTTCGTC	TACTATCTCC	GCCGCGTCTA	GTAGTCATGC	GGCGGCGACC	AGCGGTCGCA
8230	8240	8250	8260	8270	8280
CGGTGCGTGC	AGGCGCTCGT	AGAAGCTGGT	GAACCGGAGC	AAGCCGACGG	AGCCGATCAG
GCCACGCACG	TCCGCGAGCA	TCTTCGACCA	CTTGGCCTCG	TTCGGCTGCC	TCGGCTAGTC
8290	8300	8310	8320	8330	8340
CGCCACTGCG	GCGCCGAGGA	CGGTGAGCCC	GCAGACGAGA	ACGGCTGCCC	AGACGGGAAG
GCGGTGACGC	CGCGGCTCCT	GCCACTCGGG	CGTCTGCTCT	TGCCGACGGG	TCTGCCCTTC
8350	8360	8370	8380	8390	8400
GTCGGTGAGG	TGGCTCATTC	GATGATCTCC	CCGCGCATCA	GGAAGTTGCC	GAAGGCGATC
CAGCCACTCC	ACCGAGTAAG	CTACTAGAGG	GGCGCGTAGT	CCTTGAACGG	CTTCCGCTAG

FIG. 8N

8410	8420	8430	8440	8450	8460
GACGAGACGA	AGCCGATCAA	AGCCACGATC	AGGGCGGACT	CGAAATAGAG	CGAGTTGGCC
CTGCTCTGCT	TCGGCTAGTT	TCGGTGCTAG	TCCCGCCTGA	GCTTTATCTC	GCTCAACCGG
8470	8480	8490	8500	8510	8520
GTGCGGATGC	CGAAGGTCAA	GAGCATCAGC	ATGGCGTTGA	TATAGAGCGT	GTCGAGGCCG
CACGCCTACG	GCTTCCAGTT	CTCGTAGTCG	TACCGCAACT	ATATCTCGCA	CAGCTCCGGC
8530	8540	8550	8560	8570	8580
AGGATACGGT	CCTGGGCGCG	CGGTCCCCTC	ACCATGCGAT	AGAAGGCAAA	AGCCATCGCC
TCCTATGCCA	GGACCCGCGC	GCCAGGGGAG	TGGTACGCTA	TCTTCCGTTT	TCGGTAGCGG
8590	8600	8610	8620	8630	8640
AGGCCGAGCA	TGATCTGGGC	AATCAGGATC	GACCAGATGA	TTGAAAGTTC	CATCATCCGA
TCCGGCTCGT	ACTAGACCCG	TTAGTCCTAG	CTGGTCTACT	AACTTTCAAG	GTAGTAGGCT
8650	8660	8670	8680	8690	8700
ATATCTCCTT	CAGGGCGGTC	TCATAGCGCT	TGACCGTATC	GAGCCAGATG	TCCTCGTTCT
TATAGAGGAA	GTCCCGCCAG	AGTATCGCGA	ACTGGCATAG	CTCGGTCTAC	AGGAGCAAGA
8710	8720	8730	8740	8750	8760
CCATGTCGAG	CACGTGGAAG	AGCAGGGACT	TGCGGCCGCG	ATCCGGGGAA	TTC
GGTACAGCTC	GTGCACCTTC	TCGTCCCTGA	ACGCCGGCGC	TAGGCCCTT	AAG

FIG. 80

SEQUENCE: 5398-BP FRAGMENT FROM 1 TO 1200 LENGTH =1200 BP

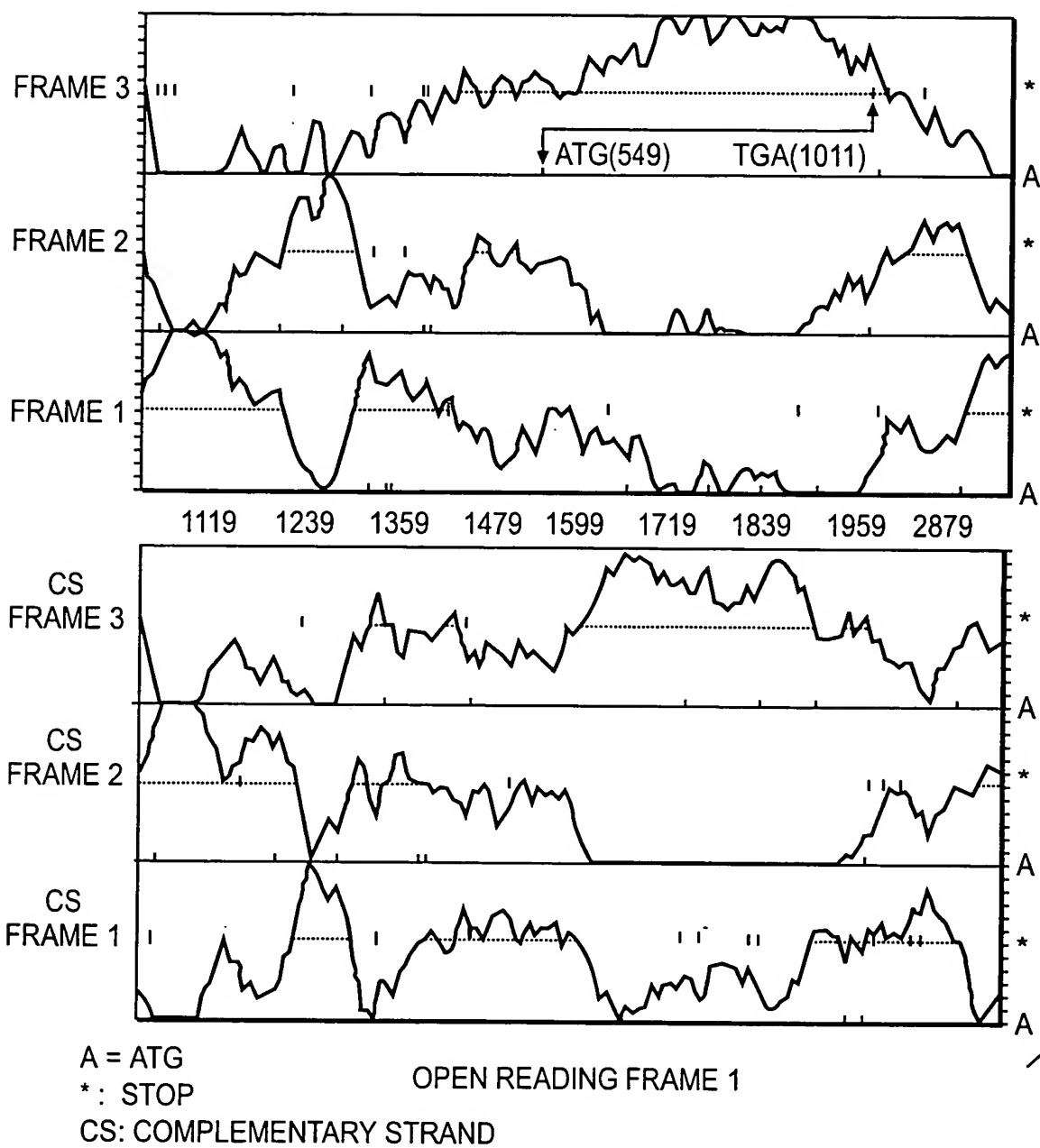


FIG. 9A

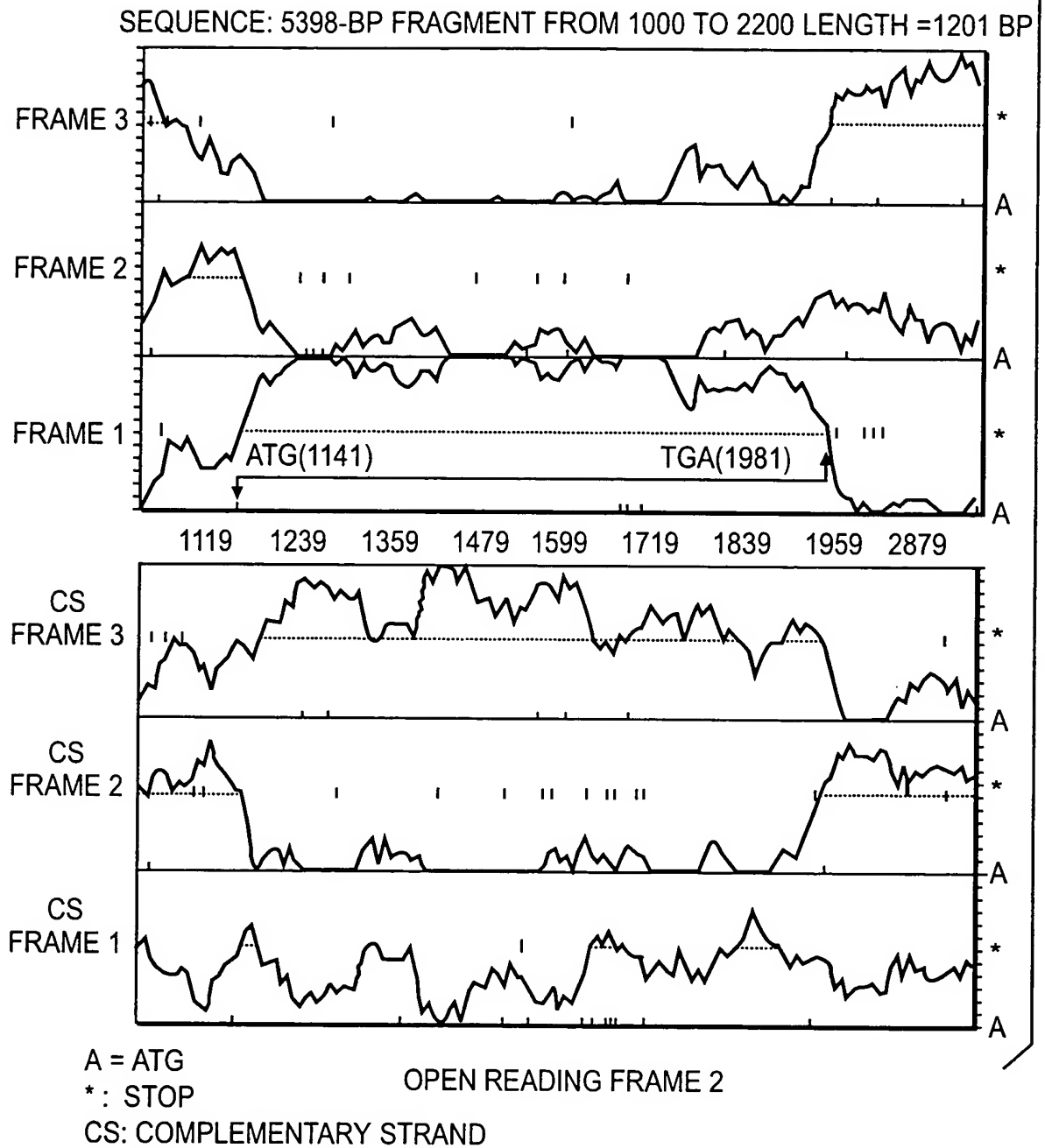


FIG. 9B

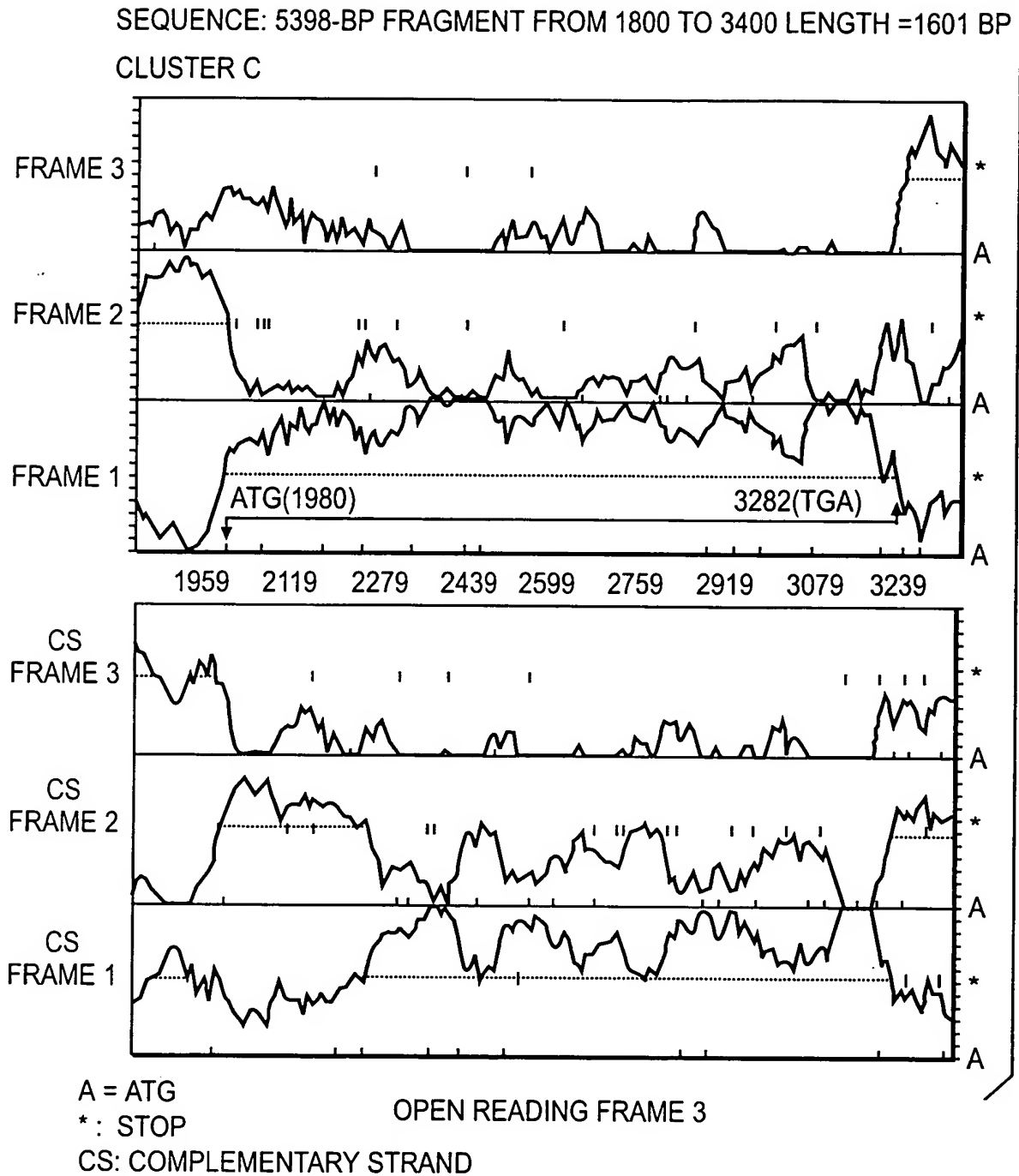


FIG. 9C

SEQUENCE: 5398-BP FRAGMENT FROM 3000 TO 4500 LENGTH =1501 BP
CLUSTER C

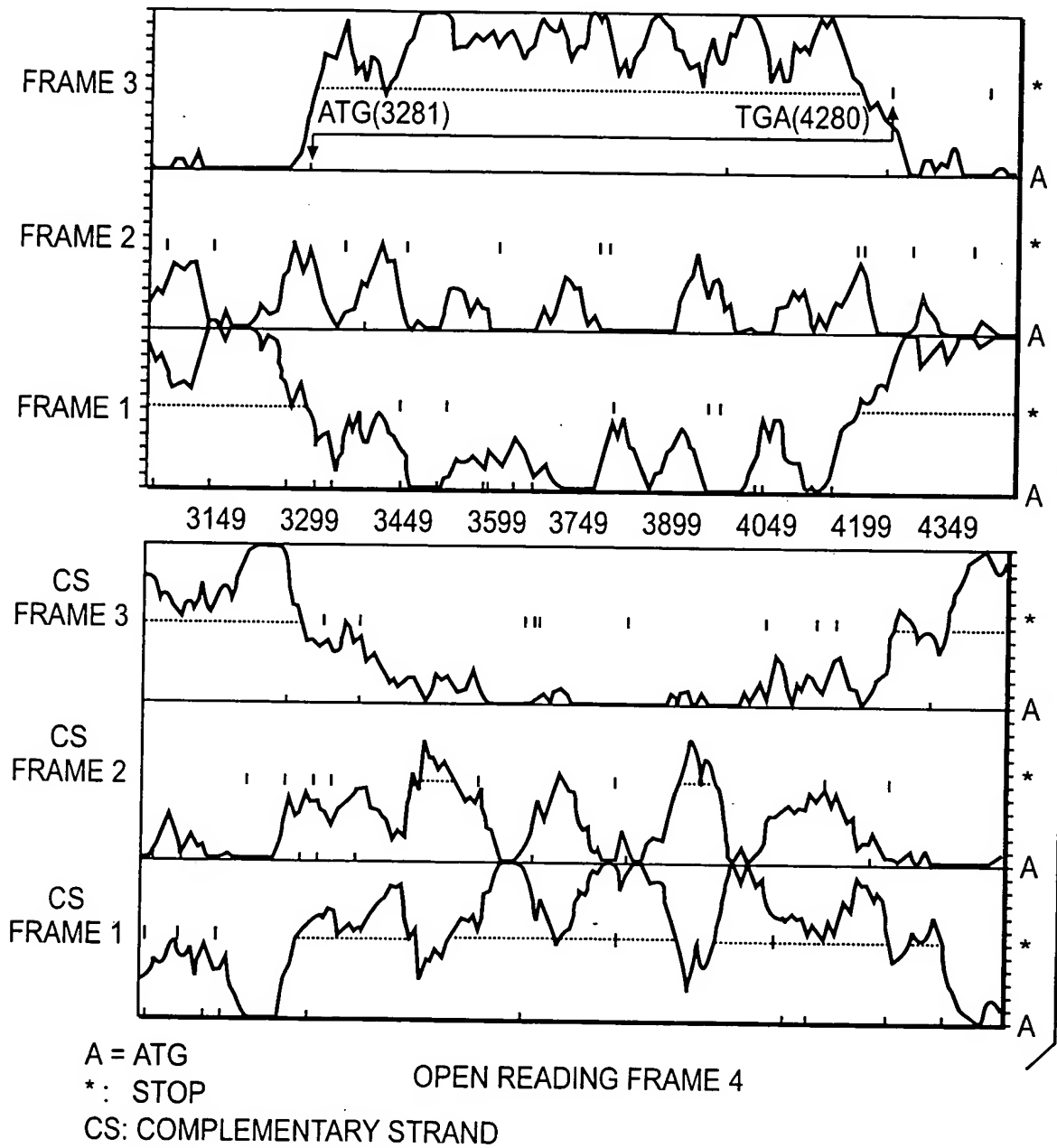


FIG. 9D

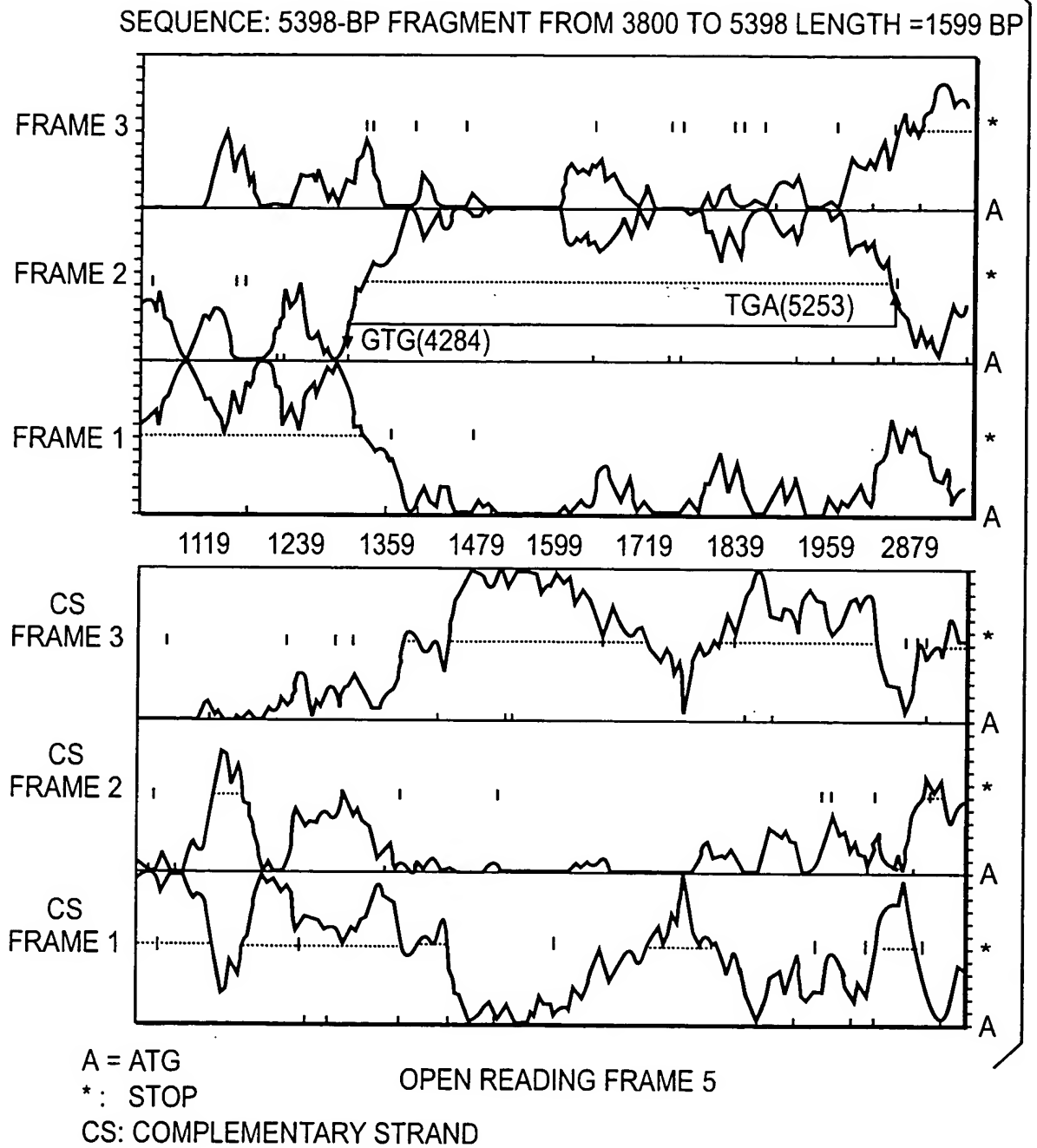
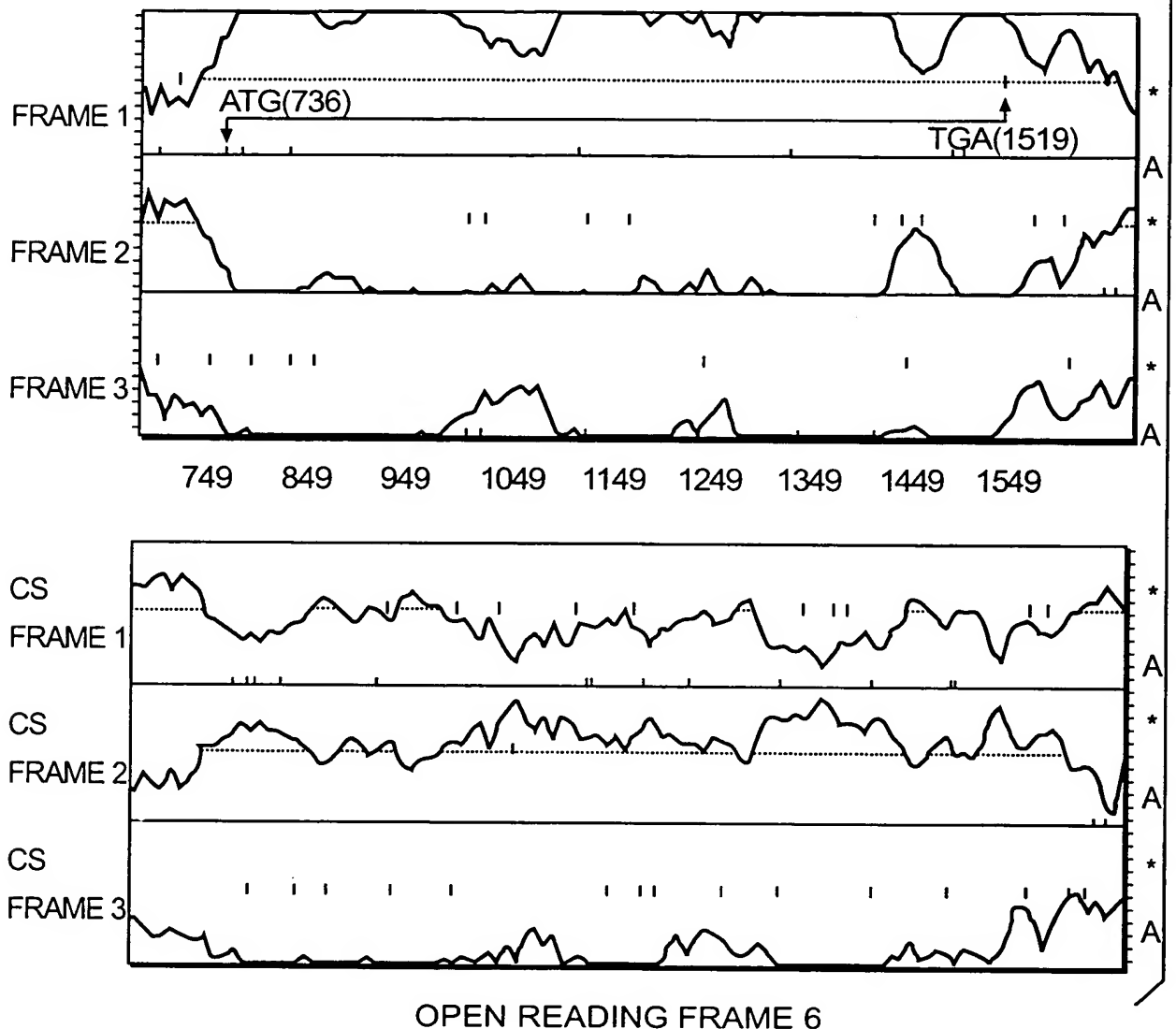


FIG. 9E

SEQUENCE: 8753-BP FRAGMENT FROM 650 TO 1650 LENGTH = 1001



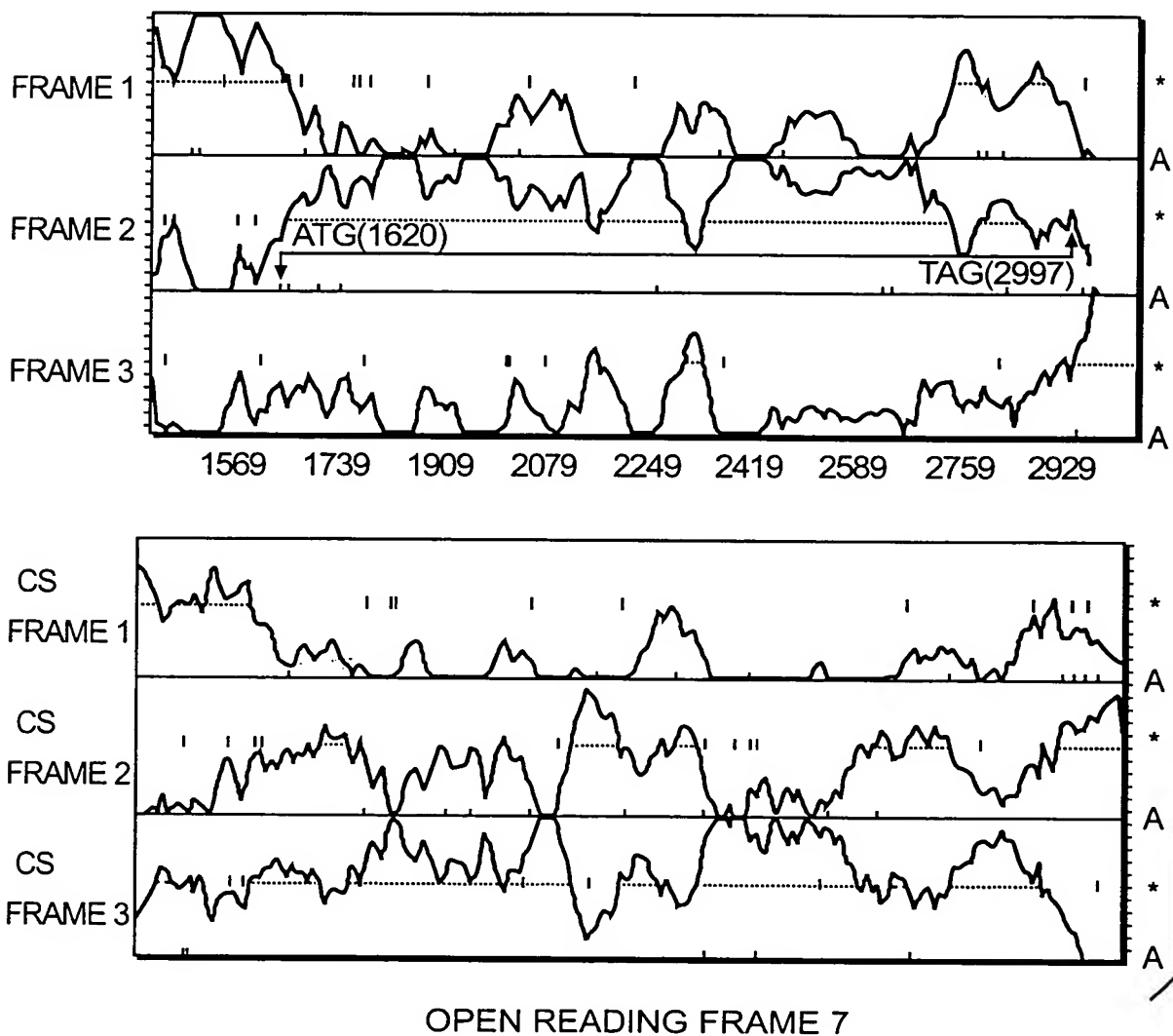
A = ATG

*: STOP

CS: COMPLEMENTARY STRAND

FIG. 10A

SEQUENCE: 8753-BP FRAGMENT FROM 1400 TO 3100 LENGTH = 1701



A = ATG

*: STOP

CS: COMPLEMENTARY STRAND

FIG. 10B

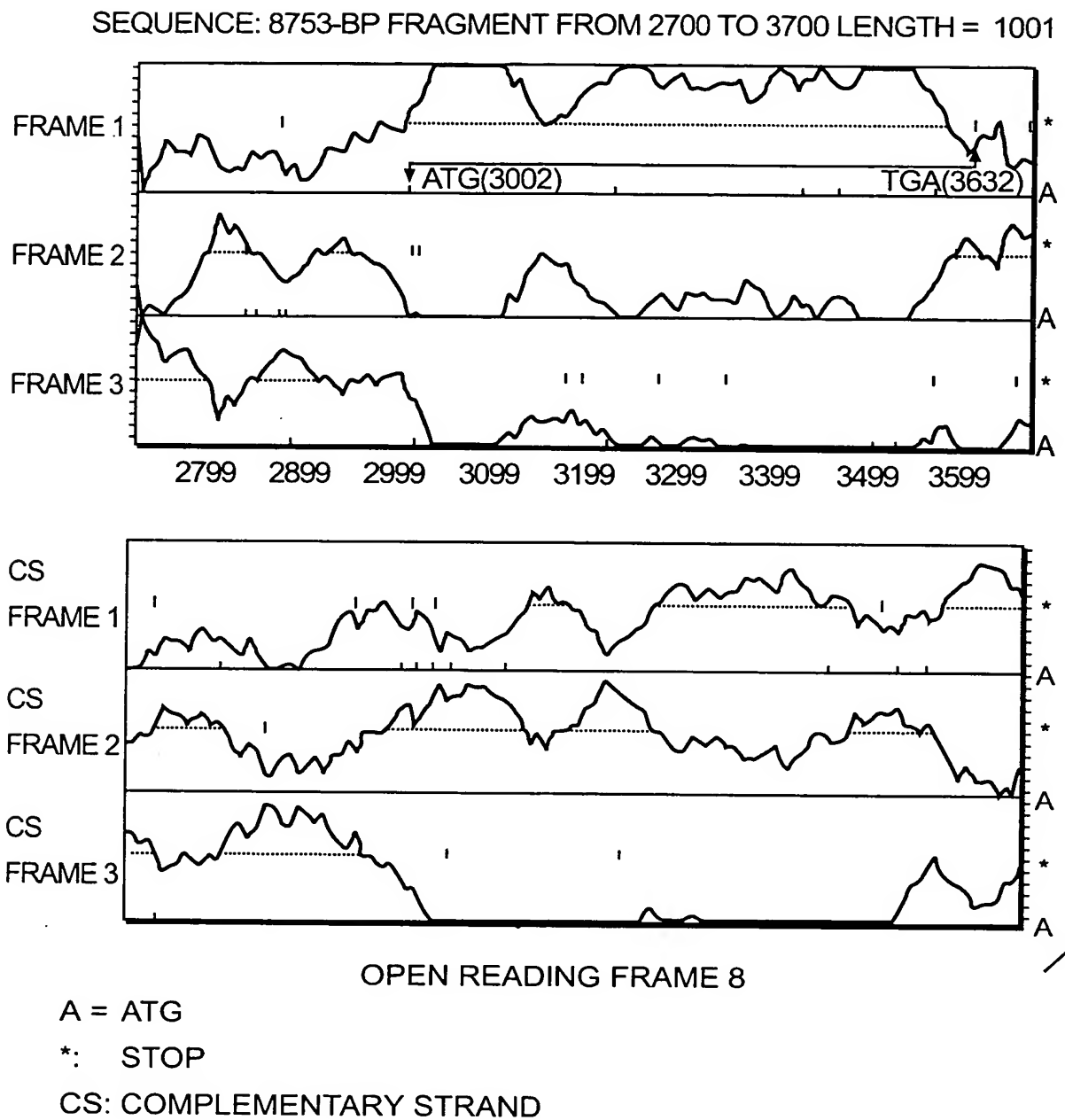


FIG. 10C

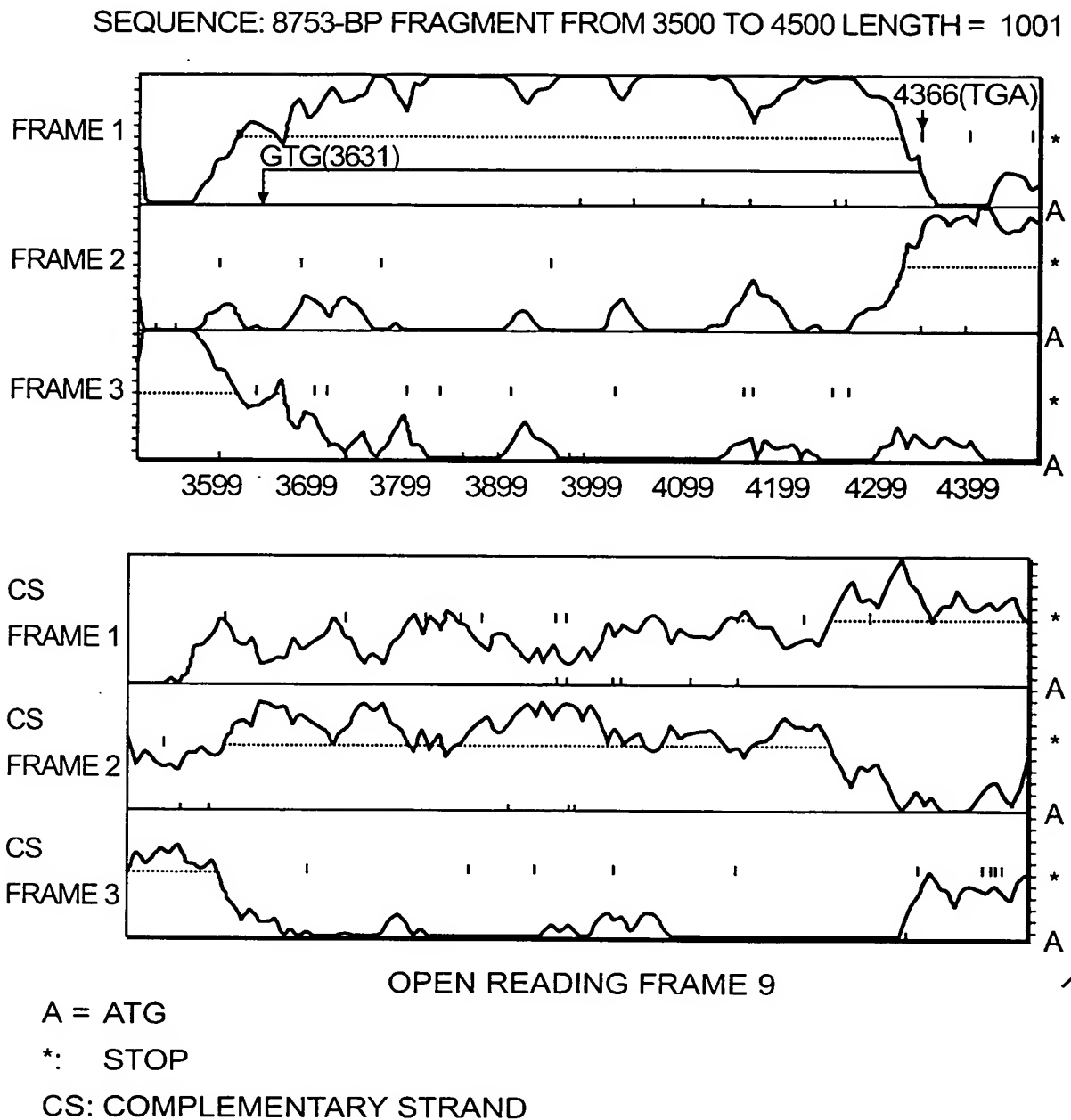


FIG. 10D

SEQUENCE: 8753-BP FRAGMENT FROM 4150 TO 5150 LENGTH = 1001

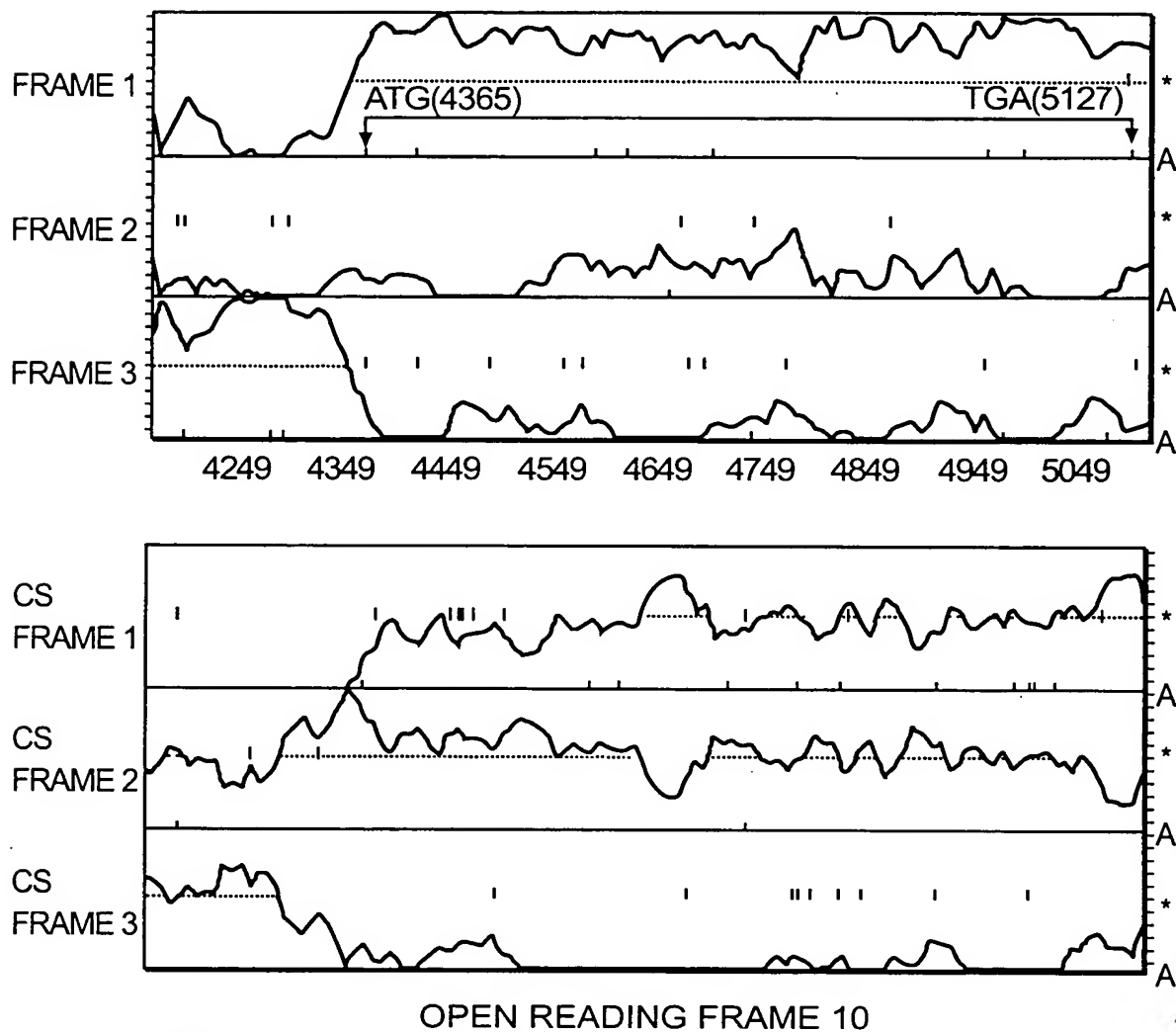


FIG. 10E

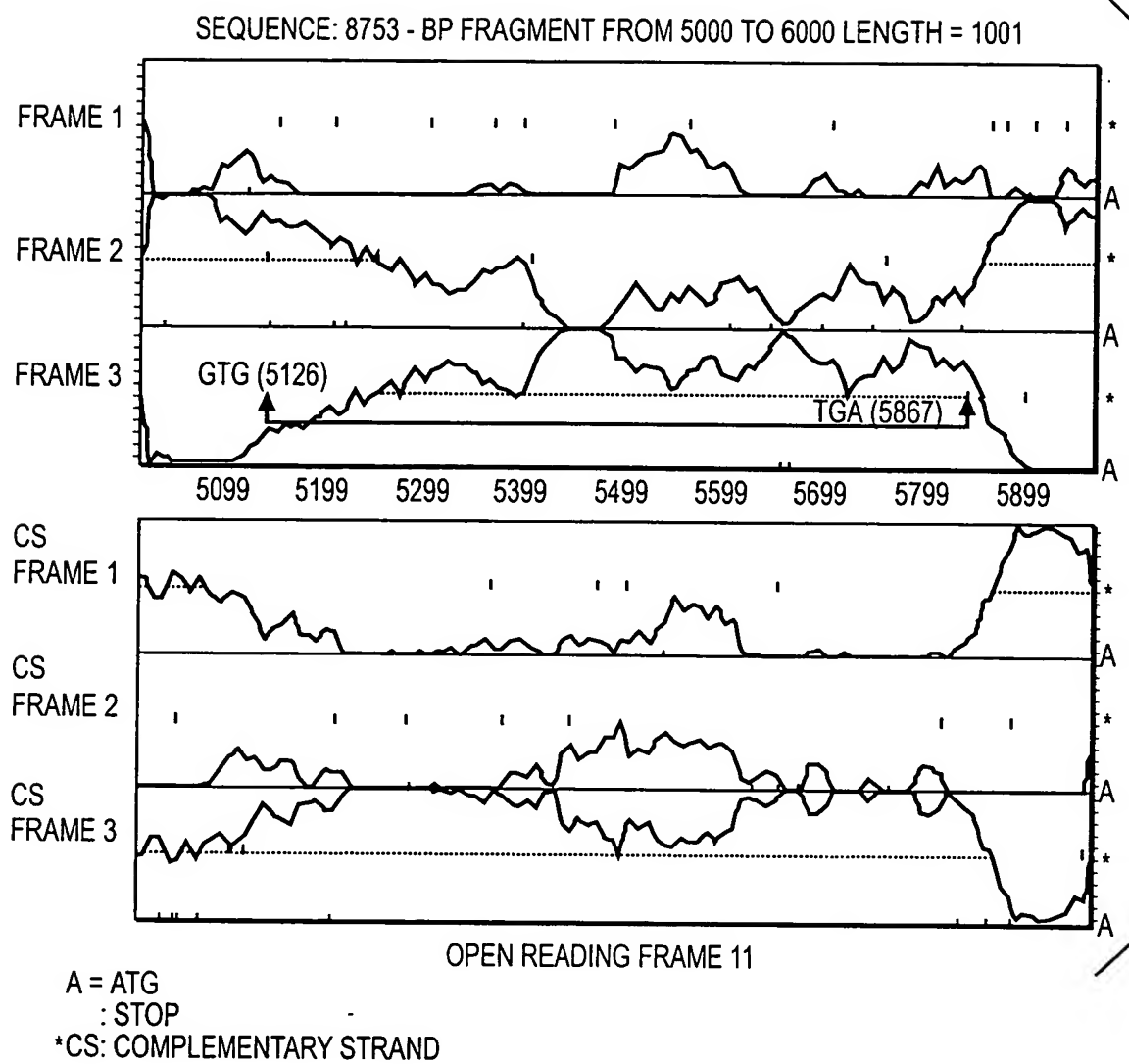


FIG. 10F

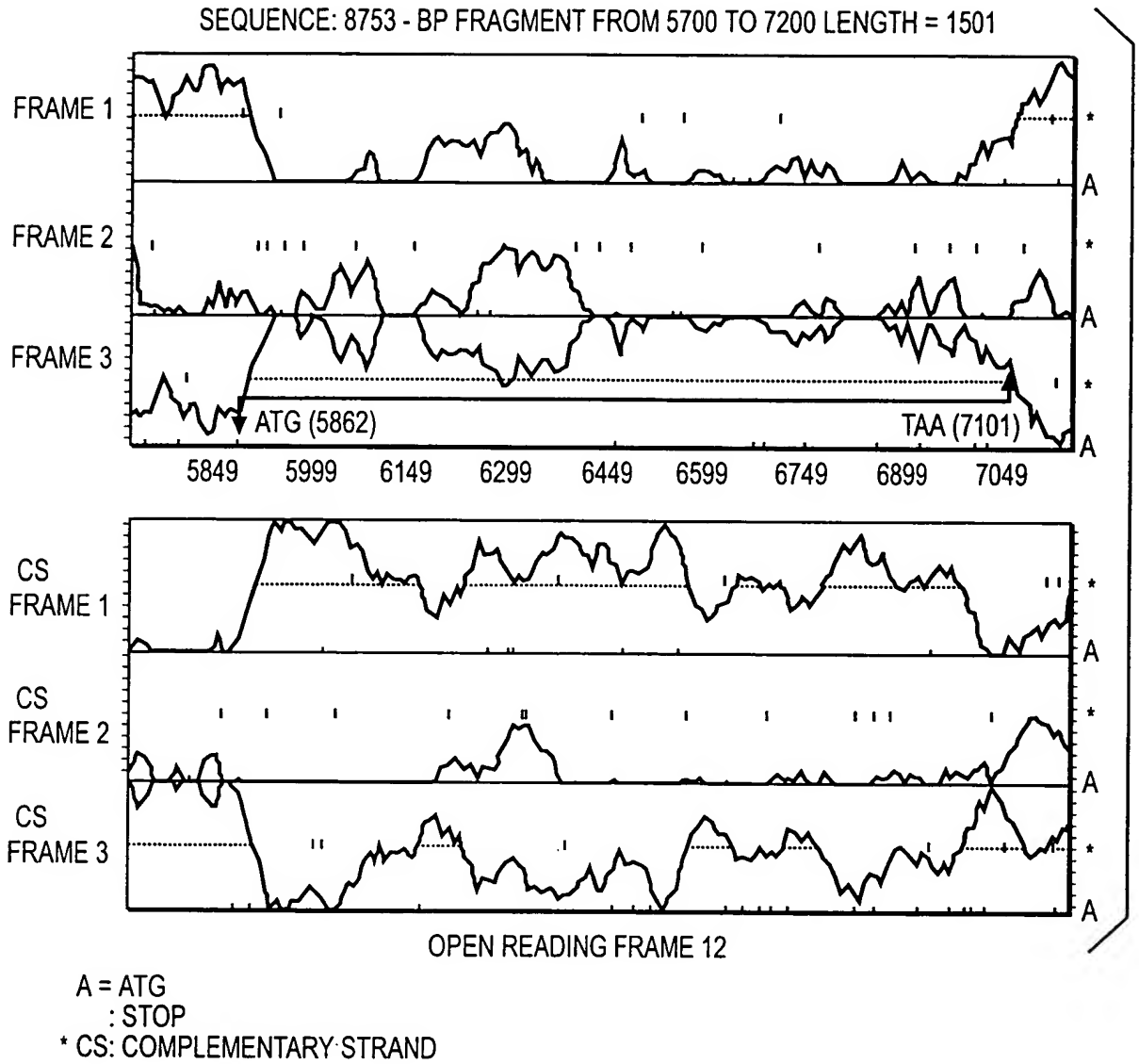


FIG. 10G

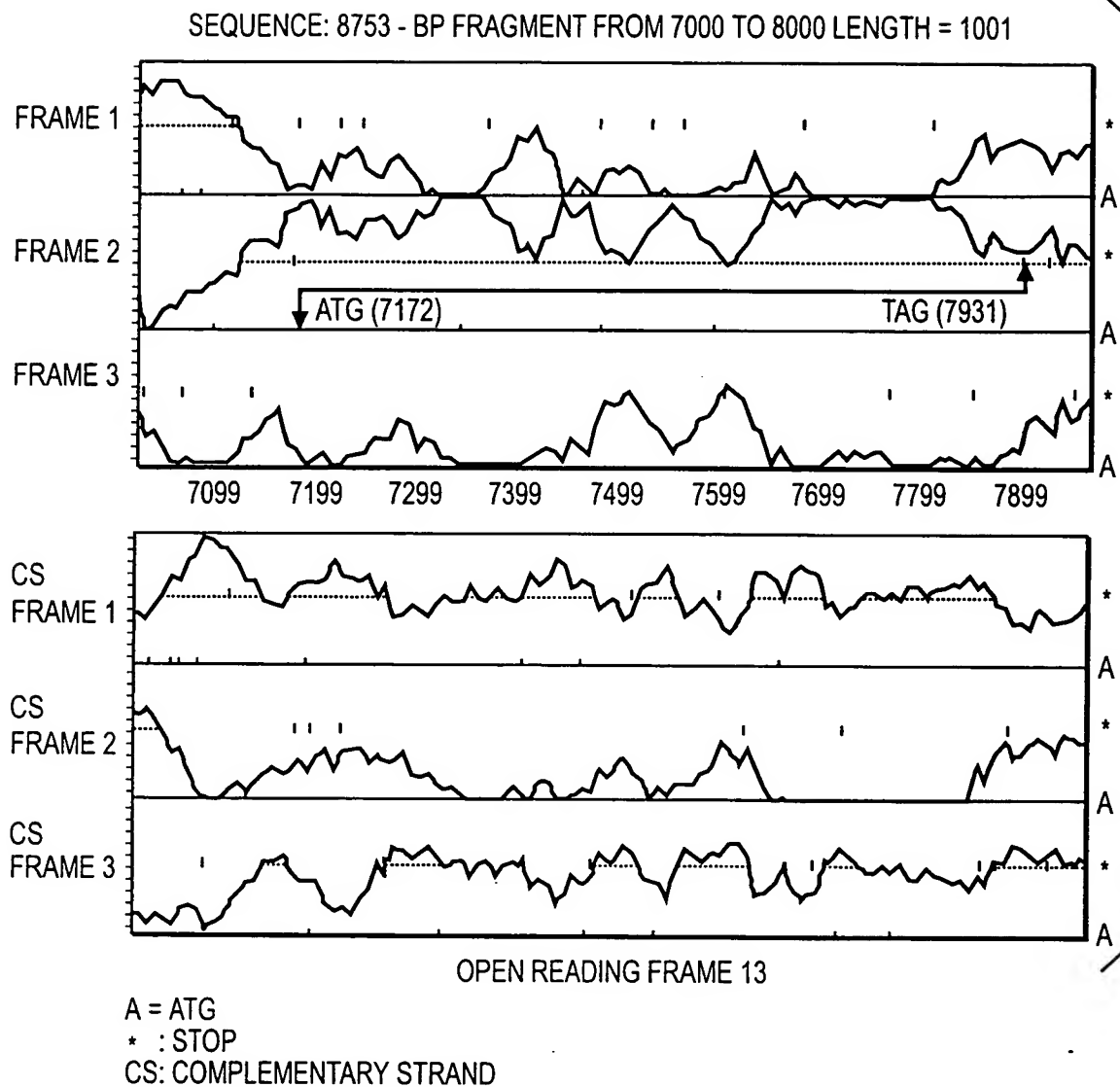
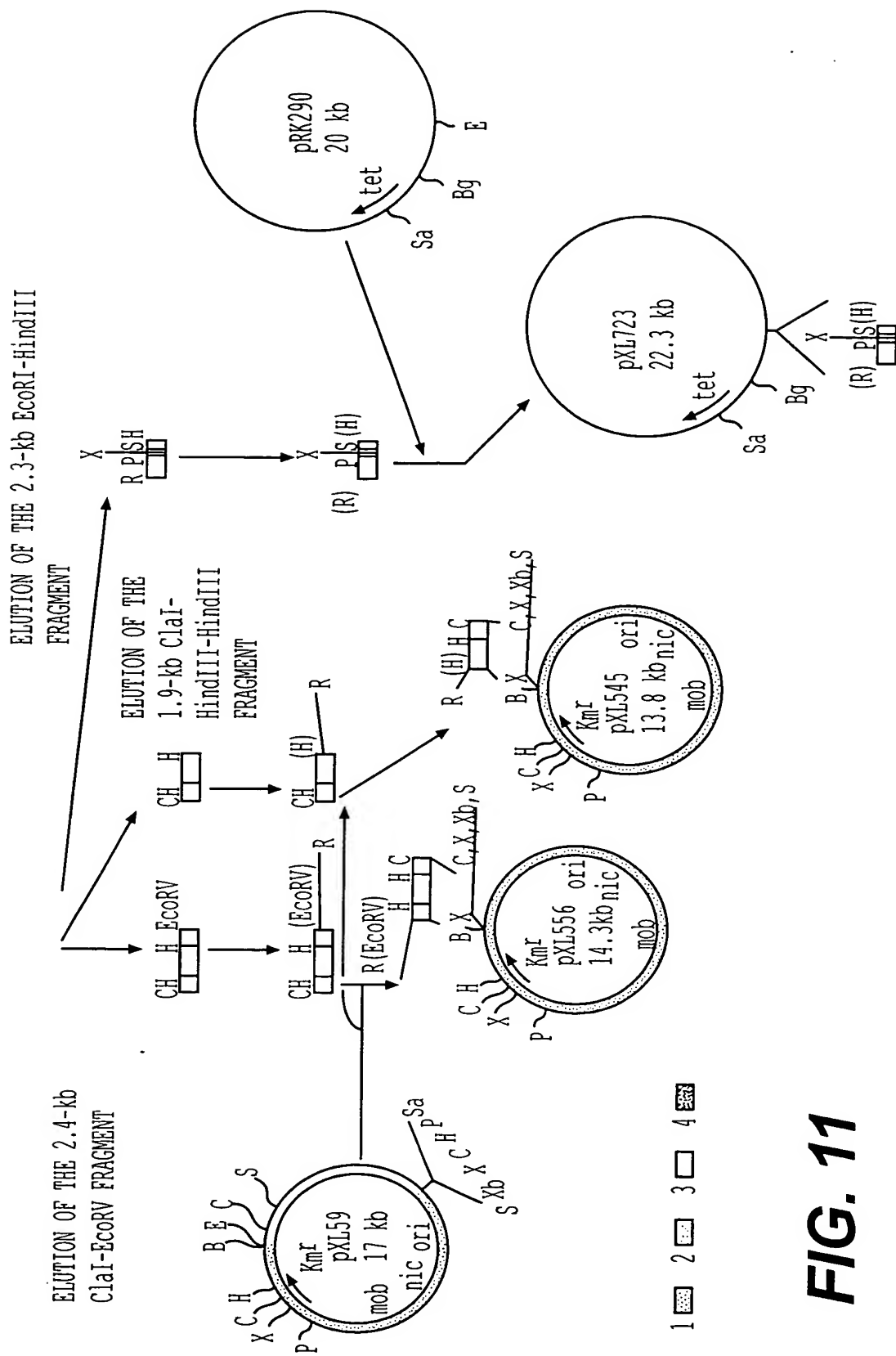
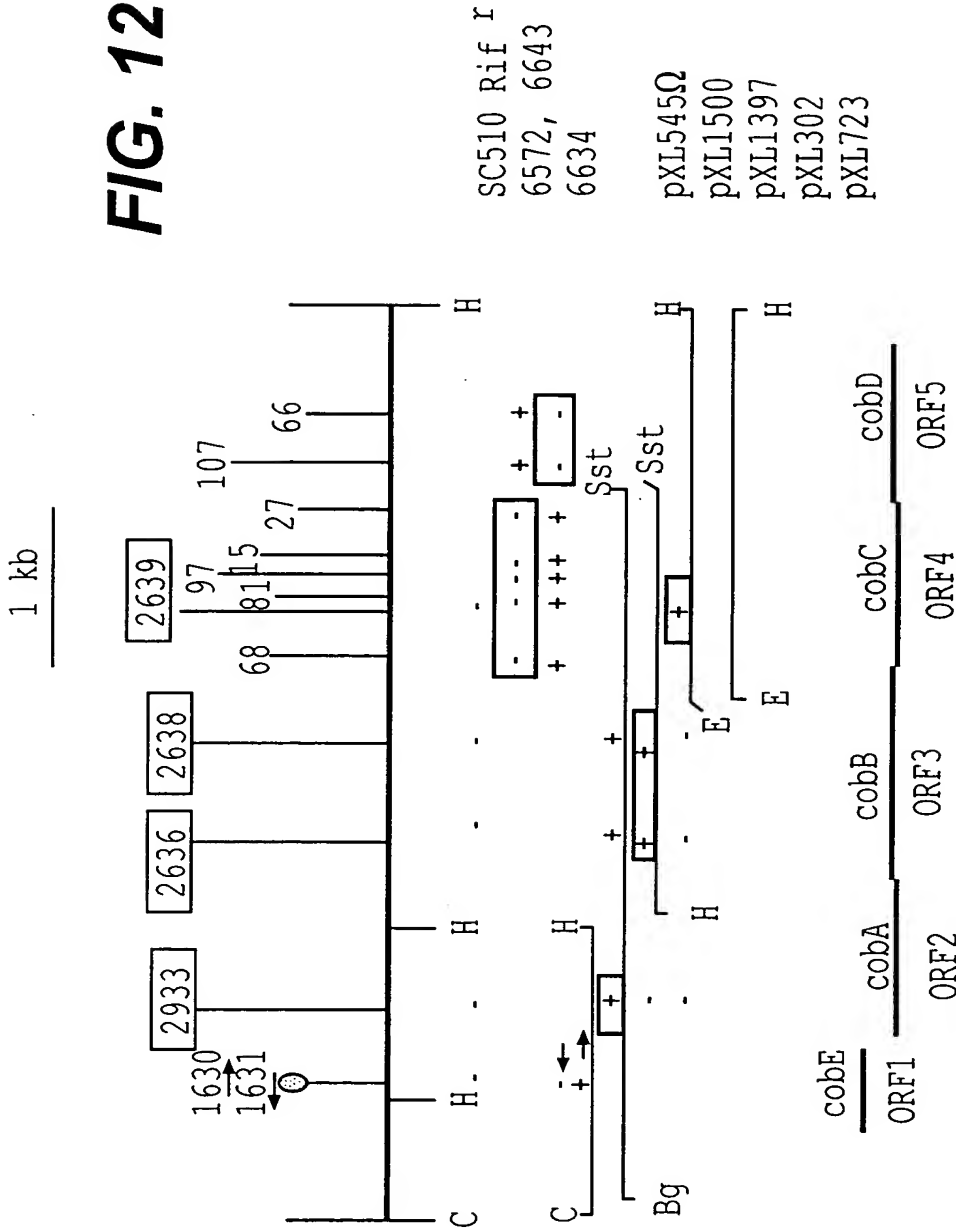


FIG. 10H



1 2 3 4

FIG. 11



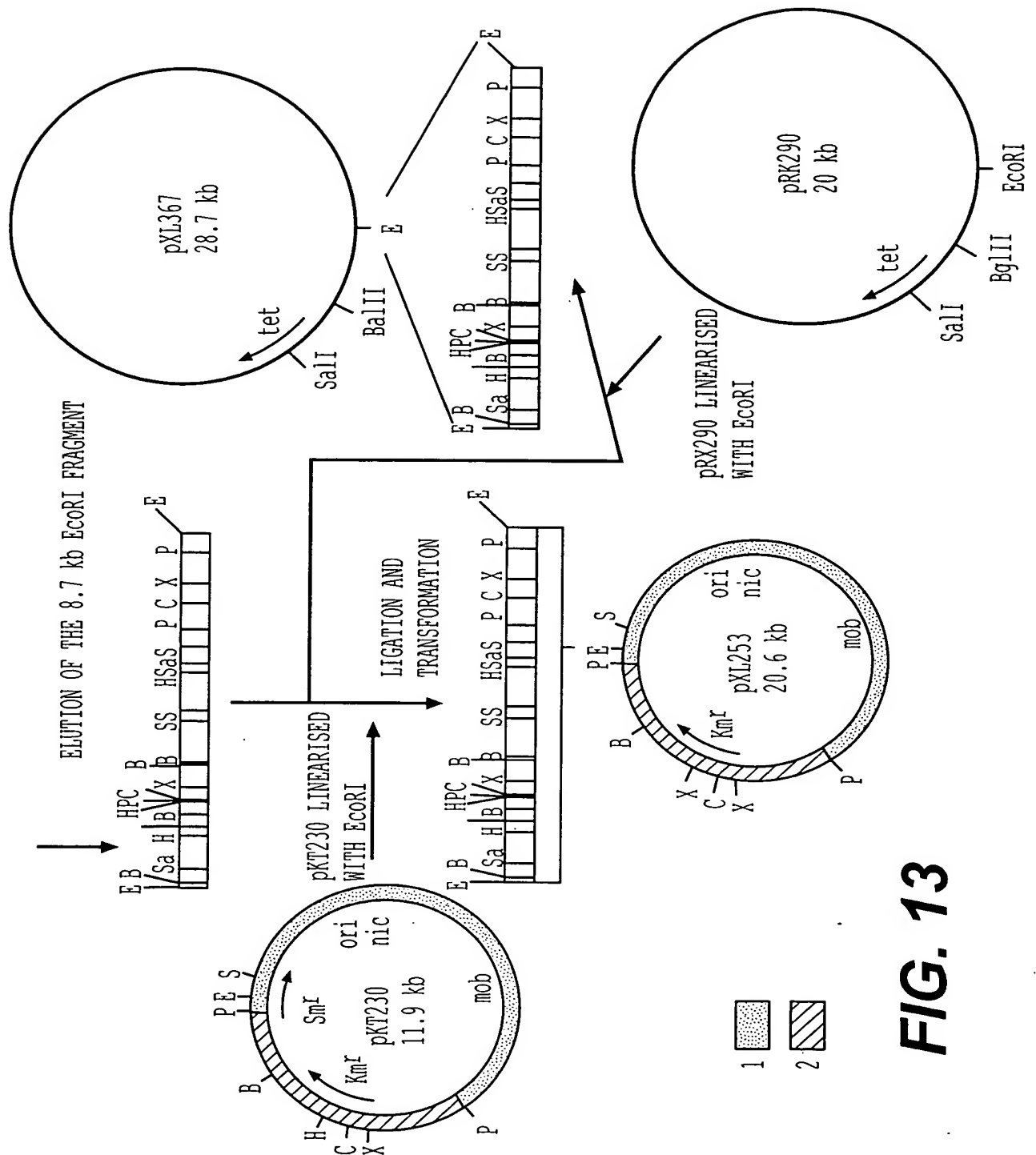


FIG. 13

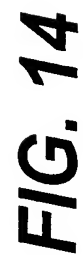


FIG. 14

cobA GENE (SEQ ID NO: 3) AND COBA PROTEIN (SEQ ID NO: 4)
 SEQUENCE OF THE 5396-BP ClaI-HindIII-HindIII-HindIII FRAGMENT
 FROM 1141 TO 1980

```

MetIleAspAspLeuPheAlaGlyLeuProAlaLeuGluLysGlySerValTrpLeuValGlyAlaGlyProGly
ATGATCGACGACCTCTTTGCCGGATTGCCGGCGCTCGAAAAAGGTTCCGGTCTGGCTGGTCCGCGCCGGCCCCGGC
  1141      1151      1161      1171      1181      1191      1201
AspProGlyLeuLeuThrLeuHisAlaAlaAsnAlaLeuArgGlnAlaAspValIleValHisAspAlaLeuVal
GATCCCCGGCCTGTTGACGCTGCATGCGGCCAATGCGCTGCGCCAGGCGGATGTGATCGTCATGATGCGCTGGTC
  1216      1226      1236      1246      1256      1266      1276
AsnGluAspCysLeuLysLeuAlaArgProGlyAlaValLeuGluPheAlaGlyLysArgGlyGlyLysProSer
AACGAGGATTGCCTGAAGCTCGCGCGGCCGGCGCCGTGCTGGAGTTTGGCGGCAAGCGTGCGCGCAAGCCGTCG
  1291      1301      1311      1321      1331      1341      1351
ProLysGlnArgAspIleSerLeuArgLeuValGluLeuAlaArgAlaGlyAsnArgValLeuArgLeuLysGly
CCGAAGCAGCGCGACATCTCGCTTCGCCTCGTCAACTCGCGCGCGCCGGCAACCGGGTGCTGCGCCTCAAAGGC
  1366      1376      1386      1396      1406      1416      1426
GlyAspProPheValPheGlyArgGlyGlyGluGluAlaLeuThrLeuValGluHisGlnValProPheArgIle
GGCGATCCCTTCGTCTTCGGTCGCGGTGGCGAGGAGGCGCTGACGTGGTCAACACCAGGTGCCGTTCCGAATC
  1441      1451      1461      1471      1481      1491      1501
ValProGlyIleThrAlaGlyIleGlyGlyLeuAlaTyrAlaGlyIleProValThrHisArgGluValAsnHis
GTGCCCCGGCATCACCGCCGGTATCGGCGGGCTTGCTATGCCGGCATTCCTGTGACCCATCGCGAGGTCAACCAC
  1516      1526      1536      1546      1556      1566      1576
AlaValThrPheLeuThrGlyHisAspSerSerGlyLeuValProAspArgIleAsnTrpGlnGlyIleAlaSer
GCGGTCACTTTCCTGACTGGCCATGATTCTCCGGCCTGGTGCCGGATCGCATCAACTGGCAGGGCATCGCCAGC
  1591      1601      1611      1621      1631      1641      1651
GlySerProValIleValMetTyrMetAlaMetLysHisIleGlyAlaIleThrAlaAsnLeuIleAlaGlyGly
GGCTCGCCTGTCATCGTCATGTACATGGCGATGAAACATATCGGCGCGATCACCGCCAACCTCATTGCCGGCGGC
  1666      1676      1686      1696      1706      1716      1726
ArgSerProAspGluProValAlaPheValCysAsnAlaAlaThrProGlnGlnAlaValLeuGluThrThrLeu
CGCTCGCCGGACGAACCGGTGCGCTTCGTCTGCAACGCCGCGACGCCGACGAGGCGGTGCTGGAAACGACGCTT
  1741      1751      1761      1771      1781      1791      1801
AlaArgAlaGluAlaAspValAlaAlaAlaGlyLeuGluProProAlaIleValValValGlyGluValValArg
GCGCGTGCAGAGGCCGATGTTGCGGCGGCAGGGCTGGAGCCGCCGCGATCGTCGTCGTCGGCGAGGTGGTGCGG
  1816      1826      1836      1846      1856      1866      1876
LeuArgAlaAlaLeuAspTrpIleGlyAlaLeuAspGlyArgLysLeuAlaAlaAspProPheAlaAsnArgIle
CTGCGCGCAGCGCTCGACTGGATCGGCGCGCTGGACGGGCGCAAGCTTGCCGCCGACCCGTTCCGCAATCGCATT
  1891      1901      1911      1921      1931      1941      1951
                                LeuArgAsnProAla***
                                CTCAGGAACCCGGCATGA
  1966      1976      1986      1996      2006      2016      2026

```

FIG. 15A

NAME = COBA

FIRST RESIDUE = 1
 LAST RESIDUE = 280

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	8	2.86	1176.56	4.02
2	LEU	L	31	11.07	3505.48	11.99
3	ILE	I	16	5.71	1809.28	6.19
4	MET	M	4	1.43	524.16	1.79
5	VAL	V	27	9.64	2674.89	9.15
6	SER	S	8	2.86	696.24	2.38
7	PRO	P	19	6.79	1843.95	6.31
8	THR	T	10	3.57	1010.50	3.46
9	ALA	A	41	14.64	2912.64	9.96
10	TYR	Y	2	0.71	326.12	1.12
11	*	*	0	0.00	0.00	0.00
12	HIS	H	7	2.50	959.42	3.28
13	GLN	Q	6	2.14	768.36	2.63
14	ASN	N	9	3.21	1026.36	3.51
15	LYS	K	8	2.86	1024.72	3.51
16	ASP	D	15	5.36	1725.45	5.90
17	GLU	E	13	4.64	1677.52	5.74
18	CYS	C	2	0.71	206.02	0.70
19	TRP	W	3	1.07	558.24	1.91
20	ARG	R	19	6.79	2965.90	10.15
21	GLY	G	32	11.43	1824.64	6.24
22	-	-	0	0.00	0.00	0.00

RESIDUES = 280
 MOLECULAR WEIGHT = 29234.
 INDEX OF POLARITY (%) = 34.
 ISOELECTRIC POINT = 7.51
 OD 260 (1mg/ml) = 0.464 OD 280 (1mg/ml) = 0.652

HYDROPHILICITY PROFILE OF THE COBA PROTEIN
 COBA FROM 1 TO 280

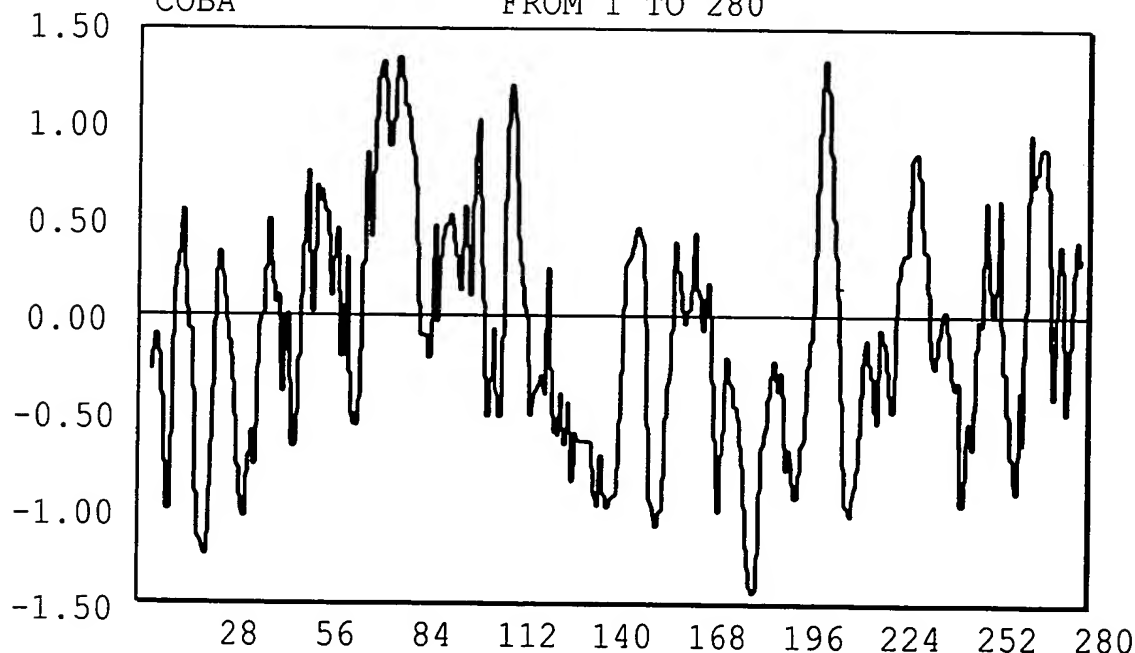


FIG. 15B

cobB GENE (SEQ ID NO: 5) AND COBB PROTEIN (SEQ ID NO: 6)
SEQUENCE OF THE 5396-BP ClaI-HindIII-HindIII-HindIII FRAGMENT
FROM 1980 TO 3281

```

MetSerGlyLeuLeuIleAlaAlaProAlaSerGlySerGlyLysThrThrValThrLeuGlyLeuMetArgAla
ATGAGCGGATTGCTGATTGCCGACCCGCGTCCGGCTCCGGCAAGACGACGGTGACGCTCGGGCTGATGCGCGCC
1980      1990      2000      2010      2020      2030      2040
LeuLysArgArgGlyValAlaIleAlaProGlyLysAlaGlyProAspTyrIleAspProAlaPheHisAlaAla
CTGAAGAGGCGCGGCGTGGCGATCGCGCCCGGAAGGCGGGCCGGACTATATCGATCCCGCTTTCACGCGGCA
2055      2065      2075      2085      2095      2105      2115
AlaThrGlyGluProCysPheAsnTyrAspProTrpAlaMetArgProGluLeuLeuLeuAlaAsnAlaSerHis
GCGACCGGCGAGCCCTGCTTCAACTACGACCCCTGGGCGATGCGCCCGAACTGCTGCTTGCCAATGCGTCGCAT
2130      2140      2150      2160      2170      2180      2190
ValAlaSerGlyGlyArgThrLeuIleValGluAlaMetMetGlyLeuHisAspGlyAlaAlaAspGlySerGly
GTGGCCTCCGGCGGGCGCACATTGATCGTCGAGGCGATGATGGGACTGCATGACGGTGCTGCCGACGGCTCGGGA
2205      2215      2225      2235      2245      2255      2265
ThrProAlaAspLeuAlaAlaThrLeuAsnLeuAlaValIleLeuValValAspCysAlaArgMetSerGlnSer
ACGCCAGCGGACCTCGCCGCGACGCTGAACCTTGCGGTCATTCTGGTGGTCGATTGCGCCCGCATGTCCCAGTCG
2280      2290      2300      2310      2320      2330      2340
ValAlaAlaLeuValArgGlyTyrAlaAspHisArgAspAspIleArgValValGlyValIleLeuAsnLysVal
GTTGCCGCCCTCGTGCGGGCTATGCGGATCATCGCGACGATATCCGGGTGGTTGGCGTCATCTCAACAAGGTC
2355      2365      2375      2385      2395      2405      2415
GlySerAspArgHisGluMetMetLeuArgAspAlaLeuGlyLysValArgMetProValPheGlyValLeuArg
GGCAGCGATCGGCATGAAATGATGCTGCGCGATGCGCTCGGCAAGGTGCGCATGCCTGTCTTCGGCGTGCTCCGG
2430      2440      2450      2460      2470      2480      2490
GlnAspSerAlaLeuGlnLeuProGluArgHisLeuGlyLeuValGlnAlaGlyGluHisSerAlaLeuGluGly
CAGGACAGCGCATTGCAACTGCCGAGCGCCATCTCGGGCTCGTGACGGCGGGCGAACACTCAGCGCTTGAGGGC
2505      2515      2525      2535      2545      2555      2565
PheIleGluAlaAlaAlaAlaArgValGluAlaAlaCysAspLeuAspAlaIleArgLeuIleAlaThrIlePhe
TTCATCGAGGCGGCGCGCGGGTCGAGGCTGCCTGCGATCTCGACGCCATCCGCTGATCGCGACGATTTTC
2580      2590      2600      2610      2620      2630      2640
ProGlnValProAlaAlaAlaAspAlaGluArgLeuArgProLeuGlyGlnArgIleAlaValAlaArgAspIle
CCGCAGGTGCCCCGCGGCGCGCGATGCCGAGCGTTTGGCGCCGCTCGGTACGCGCATCGCGTTCGCGCGGATATC
2655      2665      2675      2685      2695      2705      2715
AlaPheAlaPheCysTyrGluHisLeuLeuTyrGlyTrpArgGlnGlyGlyAlaGluIleSerPhePheSerPro
GCCTTTGCCTTCTGCTACGAGCACCTGCTTTACGGCTGGCGGCAAGGCGGCGGAGATTTCTTCTTCTCGCCG
2730      2740      2750      2760      2770      2780      2790
LeuAlaAspGluGlyProAspAlaAlaAlaAspAlaValTyrLeuProGlyGlyTyrProGluLeuHisAlaGly
CTCGCCGACGAGGGGCGGATGCGGCAGCCGATGCCGCTATCTTCCGGGGGGTTATCCGAGAGCTGCATGCGGGG
2805      2815      2825      2835      2845      2855      2865

```

FIG. 15C

GlnLeuSerAlaAlaAlaArgPheArgSerGlyMetHisSerAlaAlaGluArgGlyAlaArgIlePheGlyGlu
CAGCTGAGCGCCGCCGCCGATTCCGTTCCGGCATGCATTCCGCGGCGGAACGCGGCGCCCGCATCTTCGGCGAG
2880 2890 2900 2910 2920 2930 2940
CysGlyGlyTyrMetValLeuGlyGluGlyLeuValAlaAlaAspGlyThrArgTyrAspMetLeuGlyLeuLeu
TGCGGCGGCTATATGGTGCTCGGCGAAGGGCTTGTCGCTGCCGATGGCACACGCTACGACATGCTCGGCCTGCTG
2955 2965 2975 2985 2995 3005 3015
ProLeuValThrSerPheAlaGluArgArgArgHisLeuGlyTyrArgArgValValProValAspAsnAlaPhe
CCGCTCGTAACCAGTTTTGCCGAGCGCAGGCGGCACCTCGGCTATCGCCGCGTCGTGCCTGTCGACAACGCCTTC
3030 3040 3050 3060 3070 3080 3090
PheAspGlyProMetThrAlaHisGluPheHisTyrAlaThrIleValAlaGluGlyAlaAlaAspArgLeuPhe
TTCGATGGACCCATGACGGCGCACGAATTCCACTATGCGACCATCGTCGCCGAAGGGGCGGCCGATCGGCTGTTT
3105 3115 3125 3135 3145 3155 3165
AlaValSerAspAlaAlaGlyGluAspLeuGlyGlnAlaGlyLeuArgArgGlyProValAlaGlySerPheMet
GCGGTCAGCGACGCCGCCGCGGAGGATCTCGGCCAGGCGGGCCTCCGGCGCGGCCCTGTCGCCGGTTCCTTCATG
3180 3190 3200 3210 3220 3230 3240
HisLeuIleAspValAlaGlyAlaAla***
CATCTGATCGACGTCGCAGGTGCTGCATGA
3255 3265 3275 3285 3295 3305 3315

FIG. 15D

NAME = COBB			FIRST RESIDUE = 1		LAST RESIDUE = 434	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	17	3.92	2500.19	5.47
2	LEU	L	45	10.37	5088.60	11.14
3	ILE	I	17	3.92	1922.36	4.21
4	MET	M	14	3.23	1834.56	4.02
5	VAL	V	31	7.14	3071.17	6.72
6	SER	S	19	4.38	1653.57	3.62
7	PRO	P	21	4.84	2038.05	4.46
8	THR	T	12	2.76	1212.60	2.65
9	ALA	A	76	17.51	5399.04	11.82
10	TYR	Y	11	2.53	1793.66	3.93
11	*	*	0	0.00	0.00	0.00
12	HIS	H	14	3.23	1918.84	4.20
13	GLN	Q	9	2.07	1152.54	2.52
14	ASN	N	5	1.15	570.20	1.25
15	LYS	K	5	1.15	640.45	1.40
16	ASP	D	28	6.45	3220.84	7.05
17	GLU	E	21	4.84	2709.84	5.93
18	CYS	C	5	1.15	515.05	1.13
19	TRP	W	2	0.46	372.16	0.81
20	ARG	R	34	7.83	5307.40	11.62
21	GLY	G	48	11.06	2736.96	5.99
22	-	-	0	0.00	0.00	0.00

RESIDUES = 434
 MOLECULAR WEIGHT = 45676.
 INDEX OF POLARITY (%) = 34.
 ISOELECTRIC POINT = 6.47
 OD 260 (1mg/ml) = 0.351 OD 280 (1mg/ml) = 0.529

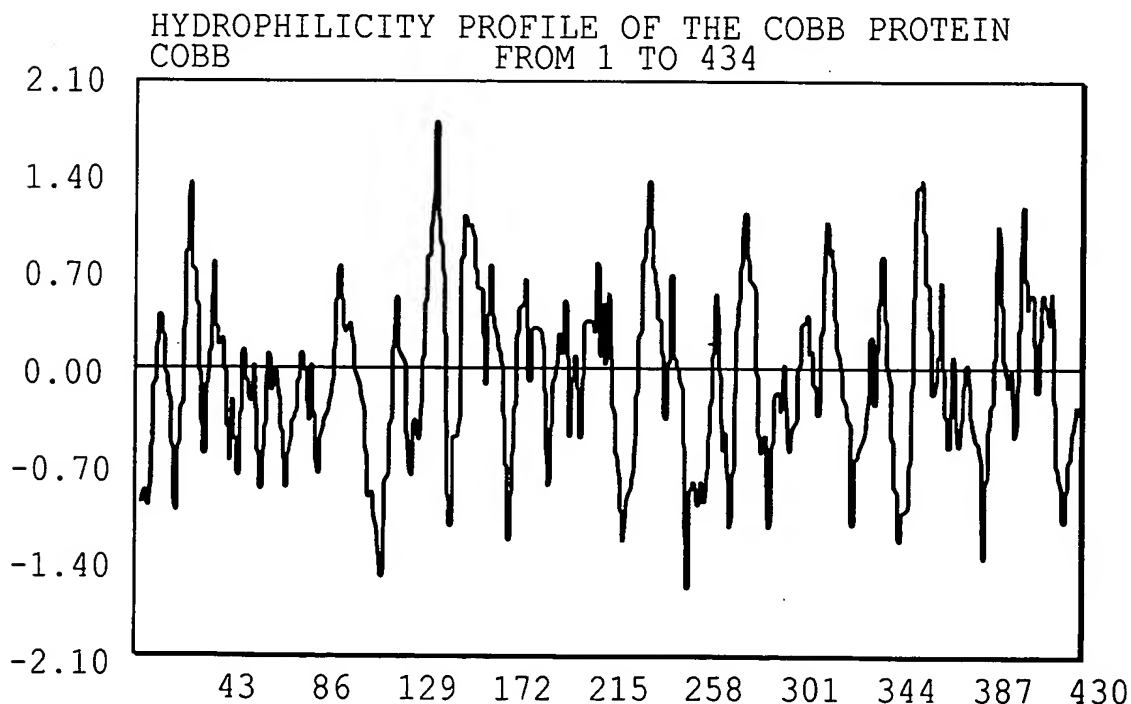


FIG. 15E

cobC GENE (SEQ ID NO: 7) AND COBC PROTEIN (SEQ ID NO: 8)
 SEQUENCE OF THE 5396-BP ClaI-HindIII-HindIII-HindIII FRAGMENT
 FROM 3281 TO 4279

```

MetSerAlaProIleValHisGlyGlyGlyIleThrGluAlaAlaAlaArgTyrGlyGlyArgProGluAspTrp
ATGAGCGCACCGATCGTTCATGGTGGCGGCATCACCAGGCCGAGCGCGCTATGGCGGCCGGCCTGAAGACTGG
3281      3291      3301      3311      3321      3331      3341
LeuAspLeuSerThrGlyIleAsnProCysProValAlaLeuProAlaValProGluArgAlaTrpHisArgLeu
CTCGATCTGTGACCGGCATCAATCCATGCCCGTCGCCTTGCCCGCGGTCCCTGAGCGCGCCTGGCACCGGCTG
3356      3366      3376      3386      3396      3406      3416
ProAspArgGlnThrValAspAspAlaArgSerAlaAlaAlaAspTyrTyrArgThrAsnGlyValLeuProLeu
CCGGATCGGCAGACGGTAGATGATGCGCGGAGCGCCGCCGCGGACTACTACCGCACCAACGGCGTGTGCCTTTG
3431      3441      3451      3461      3471      3481      3491
ProValProGlyThrGlnSerValIleGlnLeuLeuProArgLeuAlaProAlaAsnArgHisValAlaIlePhe
CCGGTGCCGGGCACCCAGTCGGTGATCCAGCTCTGCCAGTCTTGCTCCGGCCAACAGGCACGTCGCGATTTTC
3506      3516      3526      3536      3546      3556      3566
GlyProThrTyrGlyGluTyrAlaArgValLeuGluAlaAlaGlyPheAlaValAspArgValAlaAspAlaAsp
GGCGCGACCTATGGCGAGTATGCCCGCGTGCTGAAGCGGCCGGCTTTGCTGTGATCGCGTCGCGGATGCCGAC
3581      3591      3601      3611      3621      3631      3641
AlaLeuThrAlaGluHisGlyLeuValIleValValAsnProAsnAsnProThrGlyArgAlaLeuAlaProAla
GCGCTCACGGCCGAACATGGGCTTGTCATCGTCGTCACCCCAACAACCCGACCGGCCGCGCCTTGCGCCGCGC
3656      3666      3676      3686      3696      3706      3716
GluLeuLeuAlaIleAlaAlaArgGlnLysAlaSerGlyGlyLeuLeuLeuValAspGluAlaPheGlyAspLeu
GAGCTTCTGGCGATCGCCGCAAGGCAGAAGCGAGCGGCCGACTGCTGCTGGTTCGATGAGGCCTTCGGCGATCTT
3731      3741      3751      3761      3771      3781      3791
GluProGlnLeuSerValAlaGlyHisAlaSerGlyGlnGlyAsnLeuIleValPheArgSerPheGlyLysPhe
GAGCCGCAACTGAGTGTGCTGGTCACGCGTCAGGGCAAGGCAACCTCATCGTCTTCCGCTCCTTCGGCAAGTTC
3806      3816      3826      3836      3846      3856      3866
PheGlyLeuAlaGlyLeuArgLeuGlyPheValValAlaThrGluProValLeuAlaSerPheAlaAspTrpLeu
TTCGGCCTTGCGGGCCTGCGCCTCGGCTTCGTCGTTGCGACCGAGCCAGTGCTTGCATCCTTTGCCGATTGGCTC
3881      3891      3901      3911      3921      3931      3941
GlyProTrpAlaValSerGlyProAlaLeuThrIleSerLysAlaLeuMetGlnGlyAspThrLysAlaIleAla
GGTCCCTGGGCTGTCTCCGGCCCGCGTTGACGATCTCGAAAGCGCTGATGCAGGGCGATACGAAGGCGATCGCG
3956      3966      3976      3986      3996      4006      4016
AlaGlyIleLeuGluArgArgAlaGlyLeuAspAlaAlaLeuAspGlyAlaGlyLeuAsnArgIleGlyGlyThr
GCGGGCATCCTCGAGCGTCGCGCCGGCCTCGATGCGGCTCTCGATGGGGCAGGGCTCAACCGTATCGGCGGCACG
4031      4041      4051      4061      4071      4081      4091
GlyLeuPheValLeuValGluHisProArgAlaAlaLeuLeuGlnGluArgLeuCysGluAlaHisIleLeuThr
GGGCTATTCTGTGCTGGTCGAGCATCCAGGGCAGCTCTGCTGCAGGAGCGGCTCTGCGAGGCCCATATTCTCACG
4106      4116      4126      4136      4146      4156      4166
ArgLysPheAspTyrAlaProThrTrpLeuArgValGlyLeuAlaProAspAlaAlaGlyAspArgArgLeuAla
CGCAAGTTCGACTATGCCCCGACCTGGCTCAGGGTCGGTCTTGCGCCTGACGCGGCTGGTGACCGACGGCTGGCG
4181      4191      4201      4211      4221      4231      4241
AspAlaLeuAlaArgMetGluLeu***
GACGCGCTTGCCCGCATGGAGCTCTGA
4256      4266      4276      4286      4296      4306      4316

```

FIG. 15F

NAME = COBC
 FIRST RESIDUE = 1
 LAST RESIDUE = 333

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	11	3.30	1617.77	4.62
2	LEU	L	43	12.91	4862.44	13.90
3	ILE	I	13	3.90	1470.04	4.20
4	MET	M	3	0.90	393.12	1.12
5	VAL	V	24	7.21	2377.68	6.79
6	SER	S	11	3.30	957.33	2.74
7	PRO	P	23	6.91	2232.15	6.38
8	THR	T	14	4.20	1414.70	4.04
9	ALA	A	56	16.82	3978.24	11.37
10	TYR	Y	6	1.80	978.36	2.80
11	*	*	0	0.00	0.00	0.00
12	HIS	H	7	2.10	959.42	2.74
13	GLN	Q	8	2.40	1024.48	2.93
14	ASN	N	8	2.40	912.32	2.61
15	LYS	K	5	1.50	640.45	1.83
16	ASP	D	19	5.71	2185.57	6.25
17	GLU	E	15	4.50	1935.60	5.53
18	CYS	C	2	0.60	206.02	0.59
19	TRP	W	5	1.50	930.40	2.66
20	ARG	R	25	7.51	3902.50	11.15
21	GLY	G	35	10.51	1995.70	5.70
22	-	-	0	0.00	0.00	0.00

RESIDUES = 333
 MOLECULAR WEIGHT = 34992.
 INDEX OF POLARITY (%) = 34.
 ISOELECTRIC POINT = 6.72
 OD 260 (1mg/ml) = 0.670 OD 280 (1mg/ml) = 0.998

HYDROPHILICITY PROFILE OF THE COBC PROTEIN
 FROM 1 TO 333

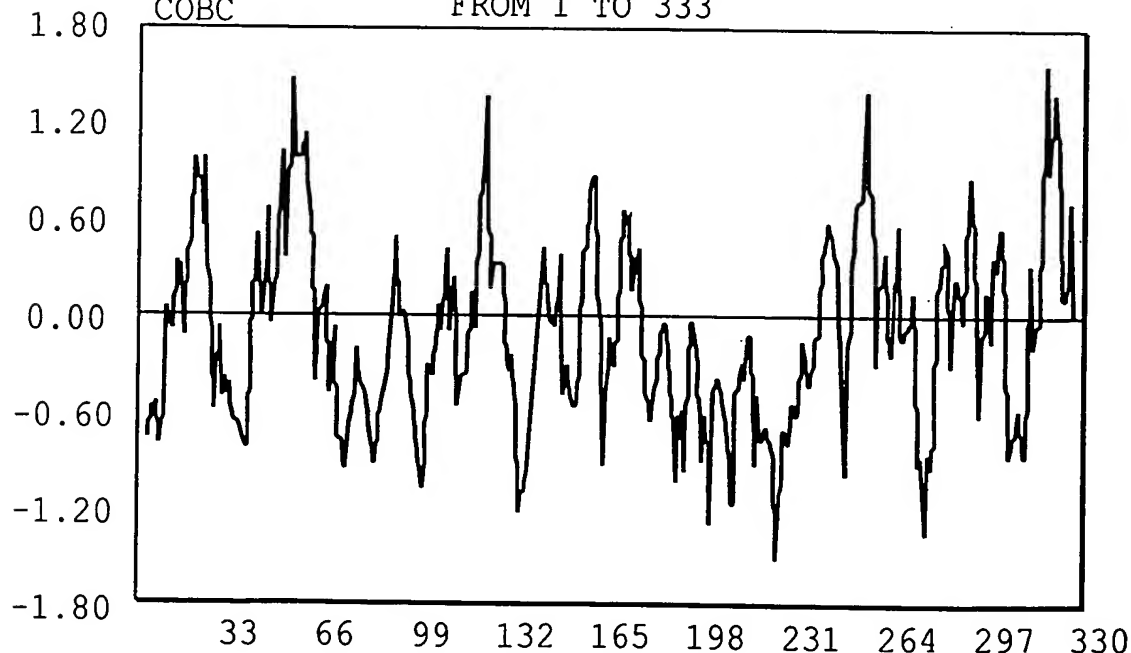


FIG. 15G

cobD GENE (SEQ ID NO: 9) AND COBD PROTEIN (SEQ ID NO: 10)
 SEQUENCE OF THE 5396-BP ClaI-HindIII-HindIII-HindIII FRAGMENT
 FROM 4284 TO 5252

```

MetSerGluThrIleLeuLeuIleLeuAlaLeuAlaLeuValIleAspArgValValGlyAspProAspTrpLeu
GTGTCGGAGACGATCCTGCTCATTCTCGCGCTGGCGCTGGTGATCGACCGCGTTGTCGGCGATCCGGACTGGCTC
4284      4294      4304      4314      4324      4334      4344
TrpAlaArgValProHisProValValPhePheGlyLysAlaIleGlyPhePheAspAlaArgLeuAsnArgGlu
TGGGCGCGCGTGCCGCATCCGGTCGTGTTTTTCGGCAAGGCCATCGGCTTTTTTCGACGCGCGGCTGAACCGGGAG
4359      4369      4379      4389      4399      4409      4419
AspLeuGluAspSerAlaArgLysPheArgGlyValValAlaIleLeuLeuLeuLeuGlyIleSerAlaTrpPhe
GACCTCGAGGATAGCGCGCGAAATTTCTGGCGTCGTCGCGATCCTTTTGTGTGCTTGGCATCAGCGCCTGGTTC
4434      4444      4454      4464      4474      4484      4494
GlyHisLeuLeuHisArgLeuPheAlaValLeuGlyProLeuGlyPheLeuLeuGluAlaValLeuValAlaVal
GGCCATCTGCTGCATCGCCTGTTCCGCGTCTCGGACCGCTCGGCTTCTGCTCGAGGCGGTTCTGGTCGCGGTC
4509      4519      4529      4539      4549      4559      4569
PheLeuAlaGlnLysSerLeuAlaAspHisValArgArgValAlaGlyGlyLeuArgGlnGlyGlyLeuGluGly
TTCCTGGCACAGAAGAGCCTCGCCGATCACGTGCGTCGCGTGGCCGGGGGCTTGCACAGGGCGGGCTGGAAGGC
4584      4594      4604      4614      4624      4634      4644
GlyArgAlaAlaValSerMetIleValGlyArgAspProLysThrLeuAspGluProAlaValCysArgAlaAla
GGGCGTGCCGCCGTGTCGATGATCGTTGGTCGCGATCCAAAGACGCTCGACGAGCCGGCGGTCTGCCGTGCCGCG
4659      4669      4679      4689      4699      4709      4719
IleGluSerLeuAlaGluAsnPheSerAspGlyValValAlaProAlaPheTrpTyrAlaValAlaGlyLeuPro
ATCGAAAGCCTTGCCGAGAATTTCTCCGACGGCGTCGTGGCGCCGGCCTTCTGGTACGCGGTTGCCGGCCTGCCG
4734      4744      4754      4764      4774      4784      4794
GlyLeuLeuAlaTyrLysMetLeuAsnThrAlaAspSerMetIleGlyHisLysSerProLysTyrLeuHisPhe
GGGCTTCTTGCCCTACAAGATGCTGAACACCGCCGATTCGATGATCGGCCACAAGTCGCCGAAATATCTGCACTTC
4809      4819      4829      4839      4849      4859      4869
GlyTrpAlaSerAlaArgLeuAspAspLeuAlaAsnLeuProAlaAlaArgLeuSerIleLeuLeuIleSerAla
GGCTGGGCCCTCGGCCGACTCGACGATCTCGCCAACCTGCCGCGAGCGAGGCTCTCGATCCTTTTGATCTCAGCC
4884      4894      4904      4914      4924      4934      4944
GlyAlaLeuIleHisArgGlyAlaSerAlaAlaLysAspAlaLeuThrValAlaLeuArgAspHisGlyLeuHis
GGTGCGCTGATCCATCGTGGCGCCAGCGCCCAAGGATGCGCTGACCGTGCCCTTCGCGACCATGGCCTGCAC
4959      4969      4979      4989      4999      5009      5019
ArgSerProAsnSerGlyTrpProGluAlaAlaMetAlaGlyAlaLeuAspLeuGlnLeuAlaGlyProArgIle
CGCTCGCCGAACCTCCGGCTGGCCGGAAGCGGCCATGGCCGGCGCGCTCGATCTGCAGCTTGCCGGTCCGCGGATC
5034      5044      5054      5064      5074      5084      5094
TyrGlyGlyValLysValSerGluProMetIleAsnGlyProGlyArgAlaValAlaThrSerGluAspIleAsp
TATGGCGCGCTCAAGGTCAGCGAACCTATGATCAACGGTCCGGGCCGAGCGGTTGCAACAAGCGAAGACATCGAC
5109      5119      5129      5139      5149      5159      5169
AlaGlyIleAlaValPheTyrGlyAlaCysThrValMetAlaGlyPheValLeuAlaIleAlaMetIle***
GCCGTATTGCTGTATTTATGGCGCCTGTACGGTCATGGCCGGGTTGTTCTTGCAATCGCAATGATTTGA
5184      5194      5204      5214      5224      5234      5244

```

FIG. 15H

NAME = COBD			FIRST RESIDUE = 1		LAST RESIDUE = 323	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	14	4.33	2058.98	6.02
2	LEU	L	45	13.93	5088.60	14.89
3	ILE	I	18	5.57	2035.44	5.96
4	MET	M	8	2.48	1048.32	3.07
5	VAL	V	27	8.36	2674.89	7.83
6	SER	S	17	5.26	1479.51	4.33
7	PRO	P	15	4.64	1455.75	4.26
8	THR	T	6	1.86	606.30	1.77
9	ALA	A	51	15.79	3623.04	10.60
10	TYR	Y	5	1.55	815.30	2.39
11	*	*	0	0.00	0.00	0.00
12	HIS	H	9	2.79	1233.54	3.61
13	GLN	Q	3	0.93	384.18	1.12
14	ASN	N	6	1.86	684.24	2.00
15	LYS	K	9	2.79	1152.81	3.37
16	ASP	D	18	5.57	2070.54	6.06
17	GLU	E	11	3.41	1419.44	4.15
18	CYS	C	2	0.62	206.02	0.60
19	TRP	W	6	1.86	1116.48	3.27
20	ARG	R	20	6.19	3122.00	9.14
21	GLY	G	33	10.22	1881.66	5.51
22	-	-	0	0.00	0.00	0.00

RESIDUES = 323
 MOLECULAR WEIGHT = 34175.
 INDEX OF POLARITY (%) = 31.
 ISOELECTRIC POINT = 8.00
 OD 260 (1mg/ml) = 0.789 OD 280 (1mg/ml) = 1.150

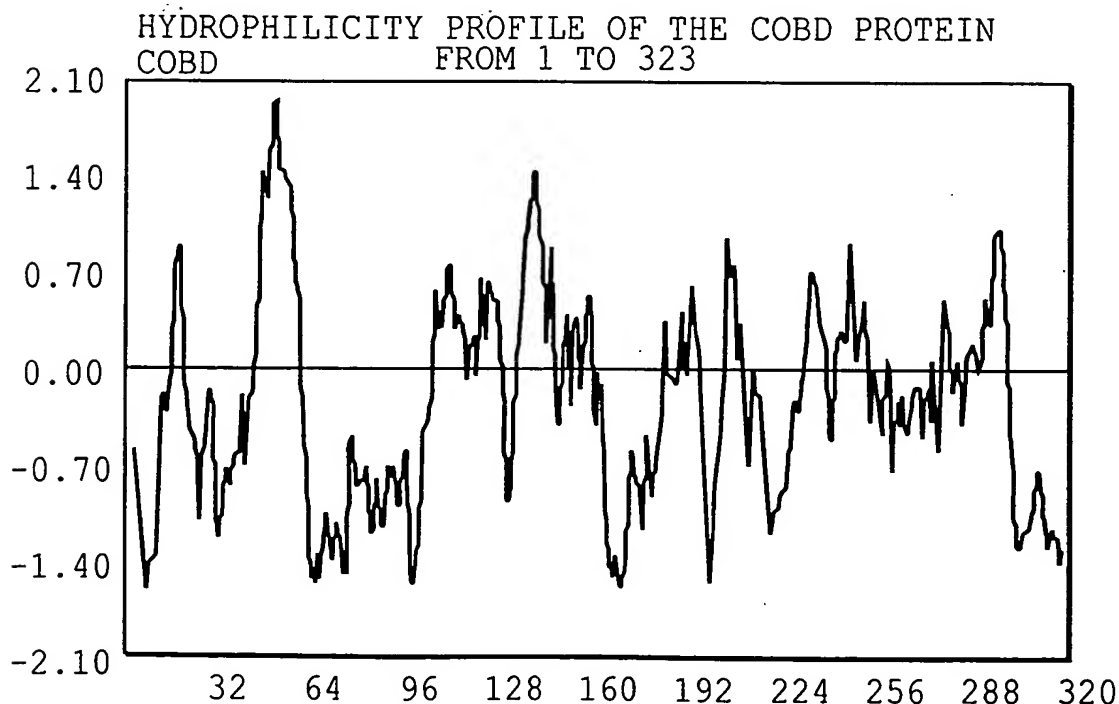


FIG. 15I

cobE GENE (SEQ ID NO: 11) AND COBE PROTEIN (SEQ ID NO: 12)
 SEQUENCE OF THE 5396-BP ClaI-HindIII-HindIII-HindIII FRAGMENT
 FROM 549 TO 1010

```

MetProSerGlyGlnHisSerAlaGlnThrThrLysAlaGlyAlaGlyLeuValLeuGlyLeuGlyCysGluArg
ATGCCATCGGGCCAACACTCTGCACAGACGACGAAAGCAGGAGCCGGGCTGGTGCTCGGGCTCGGCTGCGAGCGT
549      559      569      579      589      599      609
ArgThrProAlaGluGluValIleAlaLeuAlaGluArgAlaLeuAlaAspAlaGlyValAlaProGlyAspLeu
CGCACGCCGGCCGAAGAGGTGATCGCCCTTGCCGAGCGTGCGCTTGCCGATGCCGGTGTTGCGCCCGGCGATCTG
624      634      644      654      664      674      684
ArgLeuValAlaSerLeuAspAlaArgAlaGluGluProAlaIleLeuAlaAlaAlaGlnHisPheAlaValPro
CGGCTGGTCGCCTCGCTCGATGCTCGCGCCGAGGAGCCGGCGATCCTGGCGGCCGCTCAGCATTTGCGGGTTCCG
699      709      719      729      739      749      759
AlaAlaPheTyrAspAlaAlaThrLeuGluAlaGluAlaSerArgLeuAlaAsnProSerGluIleValPheAla
GCCGCGTTCTACGATGCCGCCACGCTCGAAGCCGAAGCTTCCCGGCTCGCCAACCCGTCCGAGATCGTCTTTGCC
774      784      794      804      814      824      834
TyrThrGlyCysHisGlyValAlaGluGlyAlaAlaLeuValGlyAlaGlyArgGluAlaValLeuIleValGln
TACACGGGTTGTCATGGCGTTGCCGAGGGTGACGCGCTCGTCGGCGCCGGTCGCGAAGCCGTGCTGATTGTGCAG
849      859      869      879      889      899      909
LysIleValSerAlaHisAlaThrAlaAlaLeuAlaGlyProAlaThrLeuArgAlaGluLysArgIleGlnAla
AAGATCGTCTCCGCCCATGCGACGGCCGCACTTGCCGGGCCGGCGACCTTGCGCGCCGAAAAGCGCATCCAGGCG
924      934      944      954      964      974      984
AlaGluAlaVal***
GCGGAGGCTGTCTGA
999      1009      1019      1029      1039      1049      1059

```

FIG. 15J

NAME = COBE

FIRST RESIDUE = 1
 LAST RESIDUE = 154

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	3	1.95	441.21	2.85
2	LEU	L	15	9.74	1696.20	10.96
3	ILE	I	6	3.90	678.48	4.38
4	MET	M	1	0.65	131.04	0.85
5	VAL	V	12	7.79	1188.84	7.68
6	SER	S	6	3.90	522.18	3.37
7	PRO	P	7	4.55	679.35	4.39
8	THR	T	7	4.55	707.35	4.57
9	ALA	A	41	26.62	2912.64	18.82
10	TYR	Y	2	1.30	326.12	2.11
11	*	*	0	0.00	0.00	0.00
12	HIS	H	4	2.60	548.24	3.54
13	GLN	Q	5	3.25	640.30	4.14
14	ASN	N	1	0.65	114.04	0.74
15	LYS	K	3	1.95	384.27	2.48
16	ASP	D	4	2.60	460.12	2.97
17	GLU	E	13	8.44	1677.52	10.84
18	CYS	C	2	1.30	206.02	1.33
19	TRP	W	0	0.00	0.00	0.00
20	ARG	R	9	5.84	1404.90	9.08
21	GLY	G	13	8.44	741.26	4.79
22	-	-	0	0.00	0.00	0.00

RESIDUES = 154
 MOLECULAR WEIGHT = 15478.
 INDEX OF POLARITY (%) = 34.
 ISOELECTRIC POINT = 5.61
 OD 260 (1mg/ml) = 0.113 OD 280 (1mg/ml) = 0.154

HYDROPHILICITY PROFILE OF THE COBE PROTEIN
 FROM 1 TO 154

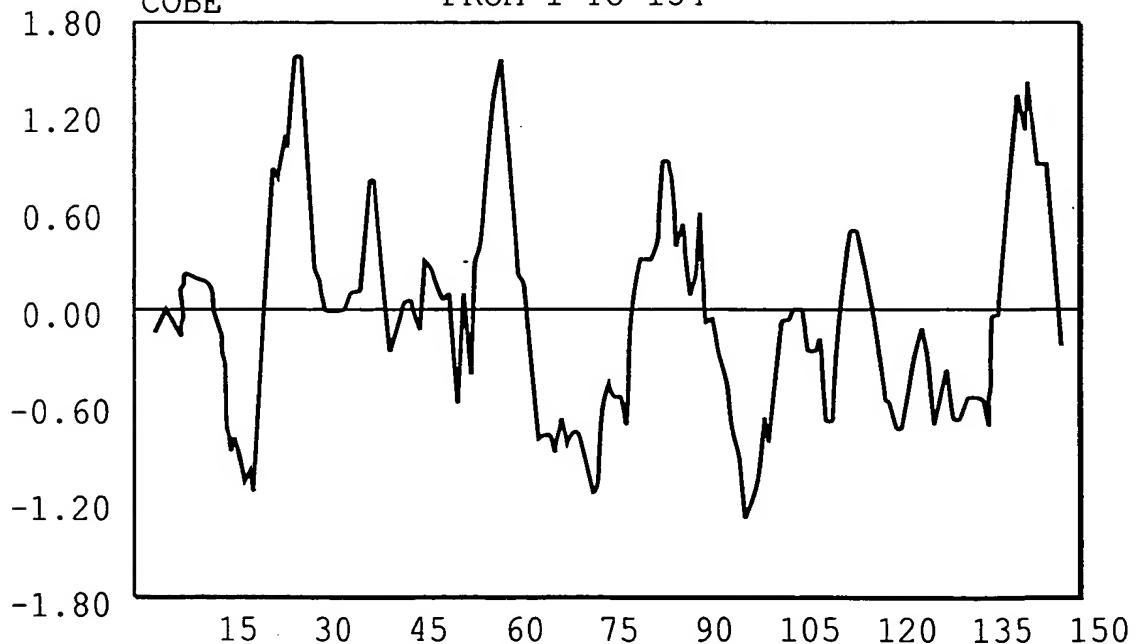


FIG. 15K

cobF GENE (SEQ ID NO: 13) AND COBF PROTEIN (SEQ ID NO: 14)
 SEQUENCE OF THE 8753-BP FRAGMENT FROM 736 TO 1521

```

MetAlaGluAlaGlyMetArgLysIleLeuIleIleGlyIleGlySerGlyAsnProGluHisMetThrValGln
ATGGCGGAGGCGGGCATGCGCAAAATTCTGATCATCGGCATCGGTTCTGGGCAATCCCGAACACATGACCGTGCAG
736       746       756       766       776       786       796
AlaIleAsnAlaLeuAsnCysAlaAspValLeuPheIleProThrLysGlyAlaLysLysThrGluLeuAlaGlu
GCGATCAACGCGCTGAACTGCGCCGACGTGCTCTTTATCCCGACCAAGGGAGCGAAGAAGACCGAGCTTGCCGAA
811       821       831       841       851       861       871
ValArgArgAspIleCysAlaArgTyrValThrArgLysAspSerArgThrValGluPheAlaValProValArg
GTGCGCCGCGACATCTGCGCCCGCTACGTCACGCGCAAGGACACCGCGACCGTCGAGTTCGCGGTGCCCCGTGCGG
886       896       906       916       926       936       946
ArgThrGluGlyValSerTyrAspGlySerValAspAspTrpHisAlaGlnIleAlaGlyIleTyrGluAlaLeu
CGCACCGAAGGCGTCAGCTATGACGGCAGCGTCGATGACTGGCACGCCCAGATCGCTGGGATTTACGAAGCGCTT
961       971       981       991      1001      1011      1021
LeuSerLysGluLeuGlyGluGluGlyThrGlyAlaPheLeuValTrpGlyAspProMetLeuTyrAspSerThr
CTATCGAAGGAGTTGGGCGAAGAGGGAAGTGGCGCGTTTCTCGTCTGGGGCGACCCGATGCTCTATGACAGCACC
1036      1046      1056      1066      1076      1086      1096
IleArgIleValGluArgValLysAlaArgGlyGluValAlaPheAlaTyrAspValIleProGlyIleThrSer
ATTCGCATCGTCGAGCGGGTCAAGGCACGCGGTGAGGTGCGCTTCGCCTACGACGTCATTCCCGGGATCACCAGT
1111      1121      1131      1141      1151      1161      1171
LeuGlnAlaLeuCysAlaSerHisArgIleProLeuAsnLeuValGlyLysProValGluIleThrThrGlyArg
CTGCAGGCGCTTTGCGCCAGCCACCGCATTCGCTGAACCTCGTCGGCAAGCCGGTGGAGATCACCACGGGGCGT
1186      1196      1206      1216      1226      1236      1246
ArgLeuHisGluSerPheProGluLysSerGlnThrSerValValMetLeuAspGlyGluGlnAlaPheGlnArg
CGGCTGCACGAAAGCTTTCCCGAGAAGAGCCAGACCTCGGTCTCATGCTCGATGGCGAACAGGCGTTTCAGCGG
1261      1271      1281      1291      1301      1311      1321
ValGluAspProGluAlaGluIleTyrTrpGlyAlaTyrLeuGlyThrArgAspGluIleValIleSerGlyArg
GTCGAGGACCCGGAGGCGGAGATCTATTGGGGCGCCTATCTCGGCACGCGGGATGAGATCGTCATTTCCGGCCGC
1336      1346      1356      1366      1376      1386      1396
ValAlaGluValLysAspArgIleLeuGluThrArgAlaAlaAlaArgAlaLysMetGlyTrpIleMetAspIle
GTGGCTGAGGTGAAGGACCGGATCCTTGAAACGCGGGCGGGCGCGCGAAGATGGGATGGATCATGGACATC
1411      1421      1431      1441      1451      1461      1471
TyrLeuLeuArgLysGlyAlaAspPheAspGlu***
TATCTCCTGCGCAAGGGCGCCGACTTCGACGAGTGA
1486      1496      1506      1516

```

FIG. 16A

COBF PROTEIN FIRST RESIDUE = 1
 LAST RESIDUE = 261

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	7	2.68	1029.49	3.56
2	LEU	L	19	7.28	2148.52	7.43
3	ILE	I	21	8.05	2374.68	8.21
4	MET	M	7	2.68	917.28	3.17
5	VAL	V	22	8.43	2179.54	7.53
6	SER	S	12	4.60	1044.36	3.61
7	PRO	P	9	3.45	873.45	3.02
8	THR	T	14	5.36	1414.70	4.89
9	ALA	A	27	10.34	1918.08	6.63
10	TYR	Y	8	3.07	1304.48	4.51
11	*	*	0	0.00	0.00	0.00
12	HIS	H	4	1.53	548.24	1.90
13	GLN	Q	6	2.30	768.36	2.66
14	ASN	N	4	1.53	456.16	1.58
15	LYS	K	12	4.60	1537.08	5.31
16	ASP	D	16	6.13	1840.48	6.36
17	GLU	E	23	8.81	2967.92	10.26
18	CYS	C	3	1.15	309.03	1.07
19	TRP	W	4	1.53	744.32	2.57
20	ARG	R	21	8.05	3278.10	11.33
21	GLY	G	22	8.43	1254.44	4.34

RESIDUES = 261
 MOLECULAR WEIGHT = 28927.
 INDEX OF POLARITY (%) = 43.
 ISOELECTRIC POINT = 5.70
 OD 260 (1mg/ml) = 0.705 OD 280 (1mg/ml) = 1.097

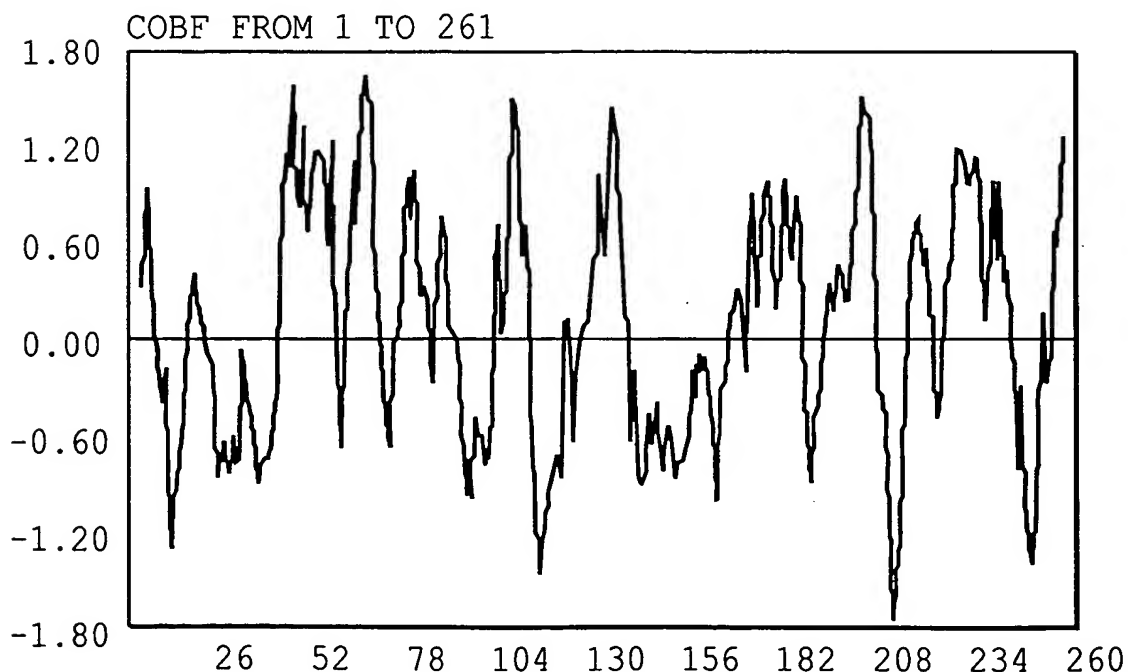


FIG. 16B

cobG GENE (SEQ ID NO: 15) AND COBG PROTEIN (SEQ ID NO: 16)
SEQUENCE OF THE 8753-BP FRAGMENT FROM 1620 TO 2999

MetThrAspLeuMetThrSerCysAlaLeuProLeuThrGlyAspAlaGlyThrValAlaSerMetArgArgGly
ATGACGGATTTGATGACCAGCTGCGCCCTTCCATTGACCGGAGATGCCGGCACCCTCGCTTCGATGCGCCGCGGC
1620 1630 1640 1650 1660 1670 1680
AlaCysProSerLeuAlaGluProMetGlnThrGlyAspGlyLeuLeuValArgValArgProThrAspAspSer
GCCTGCCCCGTCCTTGGCAGAGCCGATGCAGACCGGCGACGGCTGCTCGTGAGGGTGAGGCCAACGGATGACAGC
1695 1705 1715 1725 1735 1745 1755
LeuThrLeuProLysValIleAlaLeuAlaThrAlaAlaGluArgPheGlyAsnGlyIleIleGluIleThrAla
CTGACGCTGCCGAAGGTCATTGCCCTTGCCACGGCTGCCGAGCGCTTCGGCAATGGCATCATCGAGATTACCGCG
1770 1780 1790 1800 1810 1820 1830
ArgGlyAsnLeuGlnLeuArgGlyLeuSerAlaAlaSerValProArgLeuAlaGlnAlaIleGlyAspAlaGlu
CGCGGAAACCTGCAGCTTCGCGGCCTGAGCGCGGCTTCGGTGCCAAGGCTGGCGCAGGCGATCGGCGATGCGGAG
1845 1855 1865 1875 1885 1895 1905
IleAlaIleAlaGluGlyLeuAlaIleGluValProProLeuAlaGlyIleAspProAspGluIleAlaAspPro
ATGCCATTGCCGAGGGGCTCGCGATCGAGGTGCCGCCCTTGCCGGCATCGACCCGACGAGATCGCCGATCCG
1920 1930 1940 1950 1960 1970 1980
ArgProIleAlaThrGluLeuArgGluAlaLeuAspValArgGlnValProLeuLysLeuAlaProLysLeuSer
CGGCCGATTGCCACTGAGCTTCGTGAAGCGTTGGATGTGCGCCAGGTGCCGTTGAAGCTTGACCCAAATTATCC
1995 2005 2015 2025 2035 2045 2055
ValValIleAspSerGlyGlyArgPheGlyLeuGlyAlaValValAlaAspIleArgLeuGlnAlaValSerThr
GTCGTCATCGATAGCGGTGGCCGGTTTGGTCTCGGCGCTGTCGTCGCCGACATTGCCTTCAGGCGGTTTCGACT
2070 2080 2090 2100 2110 2120 2130
ValAlaGlyValAlaTrpValLeuSerLeuGlyGlyThrSerThrLysAlaSerSerValGlyThrLeuAlaGly
GTCGCGGGGGTGGCCTGGGTGCTGTCGCTTGCGCGCACGTCAACGAAGGCATCGAGCGTCGGGACGTTGGCCGGC
2145 2155 2165 2175 2185 2195 2205
AsnAlaValValProAlaLeuIleThrIleLeuGluLysLeuAlaSerLeuGlyThrThrMetArgGlyArgAsp
AACGCGGTCTGTCCGGCCCTGATCACCATTCTCGAGAACTGGCGAGCCTGGGCACGACGATGCGCGGGCGCGAT
2220 2230 2240 2250 2260 2270 2280
LeuAspProSerGluIleArgAlaLeuCysArgCysGluThrSerSerGluArgProAlaAlaProArgSerAla
CTGGACCCGTCGGAAATCCGCGCGCTCTGTGCTGTGAGACATCGTCCGAACGCCCGGCCGCTCCGCGTTTCGGCC
2295 2305 2315 2325 2335 2345 2355
AlaIleProGlyIleHisAlaLeuGlyAsnAlaAspThrValLeuGlyLeuGlyLeuAlaPheAlaGlnValGlu
GCAATACCCGGCATTTCATGCGCTGGGTAACGCCGACCCGTTCTCGGCCCTCGGTCTGGCCTTTGCTCAGGTGGAG
2370 2380 2390 2400 2410 2420 2430
AlaAlaAlaLeuAlaSerTyrLeuHisGlnValGlnAlaLeuGlyAlaAsnAlaIleArgLeuAlaProGlyHis
GCCGCCGCGTGGCATCCTACCTGCATCAGGTCCAGGCGCTTGGCGCCAATGCGATCCGGCTTGCGCCCGGGCAC
2445 2455 2465 2475 2485 2495 2505

FIG. 16C

AlaPhePheValLeuGlyLeuCysProGluThrAlaAlaValAlaGlnSerLeuAlaAlaSerHisGlyPheArg
GCCTTCTTCGTCCTCGGCCTTTGCCCCGAGACCGCGGCTGTGGCGCAGAGCCTGGCAGCGTCACACGGTTTTTCGC
2520 2530 2540 2550 2560 2570 2580
IleAlaGluGlnAspProArgAsnAlaIleAlaThrCysAlaGlySerLysGlyCysAlaSerAlaTrpMetGlu
ATTGCCGAGCAGGATCCGCGCAATGCGATCGCCACCTGCGCCGCGCAGCAAGGGTTGCGCCTCGGCGTGGATGGAA
2595 2605 2615 2625 2635 2645 2655
ThrLysGlyMetAlaGluArgLeuValGluThrAlaProGluLeuLeuAspGlySerLeuThrValHisLeuSer
ACCAAGGGCATGGCCGAGCGCCTCGTCGAGACGGCGCCGAATTGCTCGACGGGTCGCTCACCGTGCATCTCTCC
2670 2680 2690 2700 2710 2720 2730
GlyCysAlaLysGlyCysAlaArgProLysProSerGluLeuThrLeuValGlyAlaProSerGlyTyrGlyLeu
GGCTGCGCCAAGGGCTGCGCCCGGCCGAAGCCGTCCGAAGTACGCTTGTGCGGTGCGCCATCAGGATACGGGCTT
2745 2755 2765 2775 2785 2795 2805
ValValAsnGlyAlaAlaAsnGlyLeuProSerAlaTyrThrAspGluAsnGlyMetGlySerAlaLeuAlaArg
GTCGTAAATGGGGCTGCCAATGGCTTGCCAAGCGCTACACCGATGAGAATGGAATGGGATCCGCCCTTGCCCGG
2820 2830 2840 2850 2860 2870 2880
LeuGlyArgLeuValArgGlnAsnLysAspAlaGlyGluSerAlaGlnSerCysLeuThrArgLeuGlyAlaAla
CTCGGCCGGCTGGTGCGGCAAAACAAAGACGCTGGCGAATCGGCGCAGTCCTGTCTTACACGGCTCGGAGCTGCG
2895 2905 2915 2925 2935 2945 2955
ArgValSerAlaAlaPheGluGlnGly***
CGCGTCTCGGCAGCGTTTGAACAGGGATAG
2970 2980 2990 3000

FIG. 16D

COBG PROTEIN

FIRST RESIDUE = 1
 LAST RESIDUE = 459

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	7	1.53	1029.49	2.21
2	LEU	L	56	12.20	6332.48	13.57
3	ILE	I	21	4.58	2374.68	5.09
4	MET	M	8	1.74	1048.32	2.25
5	VAL	V	31	6.75	3071.17	6.58
6	SER	S	32	6.97	2784.96	5.97
7	PRO	P	26	5.66	2523.30	5.41
8	THR	T	27	5.88	2728.35	5.85
9	ALA	A	78	16.99	5541.12	11.88
10	TYR	Y	3	0.65	489.18	1.05
11	*	*	0	0.00	0.00	0.00
12	HIS	H	5	1.09	685.30	1.47
13	GLN	Q	13	2.83	1664.78	3.57
14	ASN	N	10	2.18	1140.40	2.44
15	LYS	K	10	2.18	1280.90	2.75
16	ASP	D	19	4.14	2185.57	4.68
17	GLU	E	24	5.23	3096.96	6.64
18	CYS	C	10	2.18	1030.10	2.21
19	TRP	W	2	0.44	372.16	0.80
20	ARG	R	29	6.32	4526.90	9.70
21	GLY	G	48	10.46	2736.96	5.87
22	-	-	0	0.00	0.00	0.00

RESIDUES = 459
 MOLECULAR WEIGHT = 46661.
 INDEX OF POLARITY (%) = 37.
 ISOELECTRIC POINT = 6.41
 OD 260 (1mg/ml) = 0.215 OD 280 (1mg/ml) = 0.315

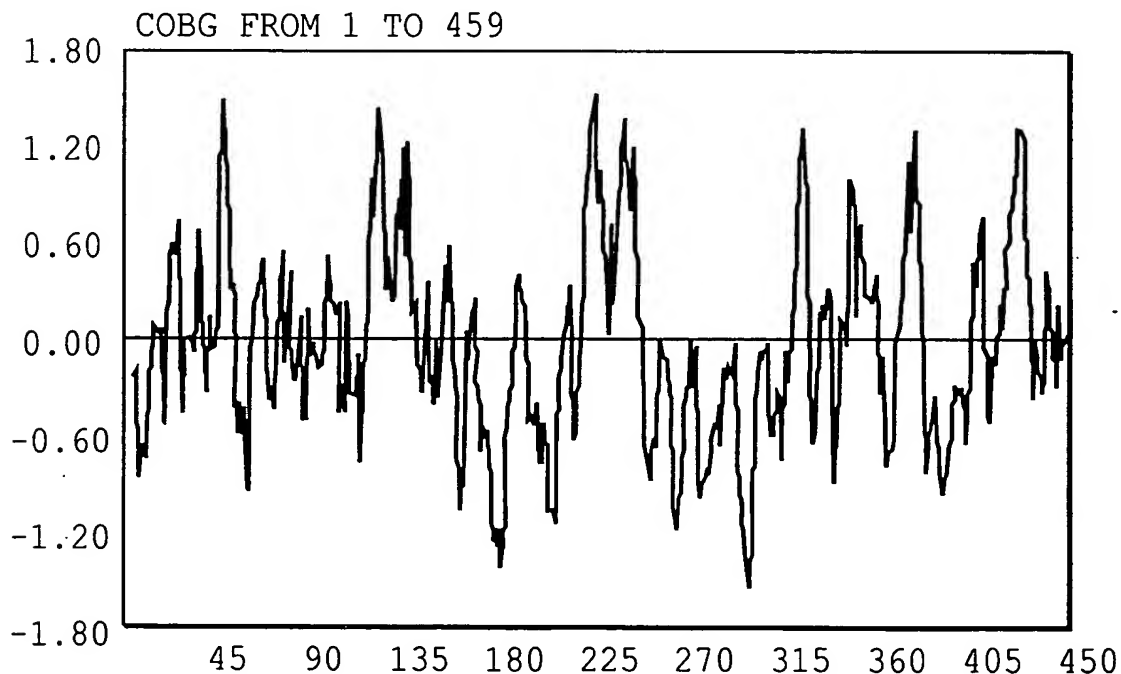


FIG. 16E

cobH GENE (SEQ ID NO: 17) AND COBH PROTEIN (SEQ ID NO: 18)
SEQUENCE OF THE 8753-BP FRAGMENT FROM 3002 TO 3634

MetProGluTyrAspTyrIleArgAspGlyAsnAlaIleTyrGluArgSerPheAlaIleIleArgAlaGluAla
ATGCCTGAGTATGATTACATTCGCGATGGCAACGCCATCTACGAGCGTTCCTTCGCCATCATCCGCGCCGAGGCC
3002 3012 3022 3032 3042 3052 3062
AspLeuSerArgPheSerGluGluGluAlaAspLeuAlaValArgMetValHisAlaCysGlySerValGluAla
GATCTGTGCGCCTTCTCCGAAGAGGAAGCGGATCTGGCTGTGCGCATGGTGACGCCTGCGGTTCCGTCGAGGCG
3077 3087 3097 3107 3117 3127 3137
ThrArgGlnPheValPheSerProAspPheValSerSerAlaArgAlaAlaLeuLysAlaGlyAlaProIleLeu
ACCAGGCAGTTCGTGTTTCTCCCGATTTCTGAAGCTCGGCGCGTGGCGGCTGAAAGCCGGTGCGCCGATCCTC
3152 3162 3172 3182 3192 3202 3212
CysAspAlaGluMetValAlaHisGlyValThrArgAlaArgLeuProAlaGlyAsnGluValIleCysThrLeu
TGCGATGCCGAGATGGTTGCGCACGGTGTACCCGCGCCCTCTGCCGGCCGGCAACGAGGTGATCTGCACGCTG
3227 3237 3247 3257 3267 3277 3287
ArgAspProArgThrProAlaLeuAlaAlaGluIleGlyAsnThrArgSerAlaAlaAlaLeuLysLeuTrpSer
CGCGATCCTCGCACGCCCGCACTTGCGGCCGAGATCGGCAACACCCGCTCCGCCGAGCCCTGAAGCTCTGGAGC
3302 3312 3322 3332 3342 3352 3362
GluArgLeuAlaGlySerValValAlaIleGlyAsnAlaProThrAlaLeuPhePheLeuLeuGluMetLeuArg
GAGCGGCTGGCCGGTTCGGTGGTCGCGATCGGCAACGCGCCGACGGCGTTGTTCTTCTTCTTGGAAATGCTGCGC
3377 3387 3397 3407 3417 3427 3437
AspGlyAlaProLysProAlaAlaIleLeuGlyMetProValGlyPheValGlyAlaAlaGluSerLysAspAla
GACGGCGCGCCGAAGCCGGCGGCAATCCTCGGCATGCCCGTCCGTTTCGTGCGGTGCGGCGGAATCGAAGGATGCG
3452 3462 3472 3482 3492 3502 3512
LeuAlaGluAsnSerTyrGlyValProPheAlaIleValArgGlyArgLeuGlyGlySerAlaMetThrAlaAla
CTGGCCGAGAACTCCTATGGCGTTCCTTCGCCATCGTGCGCGGCCGCTCGGCGGGAGTGCCATGACGGCGGCA
3527 3537 3547 3557 3567 3577 3587
AlaLeuAsnSerLeuAlaArgProGlyLeu***
GCGCTTAACCTCGCTCGCGAGGCCGGGCCTGTGA
3602 3612 3622 3632

FIG. 16F

COBH PROTEIN

FIRST RESIDUE = 1
 LAST RESIDUE = 210

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	9	4.29	1323.63	6.00
2	LEU	L	20	9.52	2261.60	10.26
3	ILE	I	10	4.76	1130.80	5.13
4	MET	M	6	2.86	786.24	3.57
5	VAL	V	14	6.67	1386.98	6.29
6	SER	S	14	6.67	1218.42	5.53
7	PRO	P	12	5.71	1164.60	5.28
8	THR	T	7	3.33	707.35	3.21
9	ALA	A	40	19.05	2841.60	12.89
10	TYR	Y	4	1.90	652.24	2.96
11	*	*	0	0.00	0.00	0.00
12	HIS	H	2	0.95	274.12	1.24
13	GLN	Q	1	0.48	128.06	0.58
14	ASN	N	6	2.86	684.24	3.10
15	LYS	K	4	1.90	512.36	2.32
16	ASP	D	9	4.29	1035.27	4.70
17	GLU	E	14	6.67	1806.56	8.19
18	CYS	C	3	1.43	309.03	1.40
19	TRP	W	1	0.48	186.08	0.84
20	ARG	R	17	8.10	2653.70	12.03
21	GLY	G	17	8.10	969.34	4.40
22	-	-	0	0.00	0.00	0.00

RESIDUES = 210
 MOLECULAR WEIGHT = 22050.
 INDEX OF POLARITY (%) = 35.
 ISOELECTRIC POINT = 6.22
 OD 260 (1mg/ml) = 0.291 OD 280 (1mg/ml) = 0.467

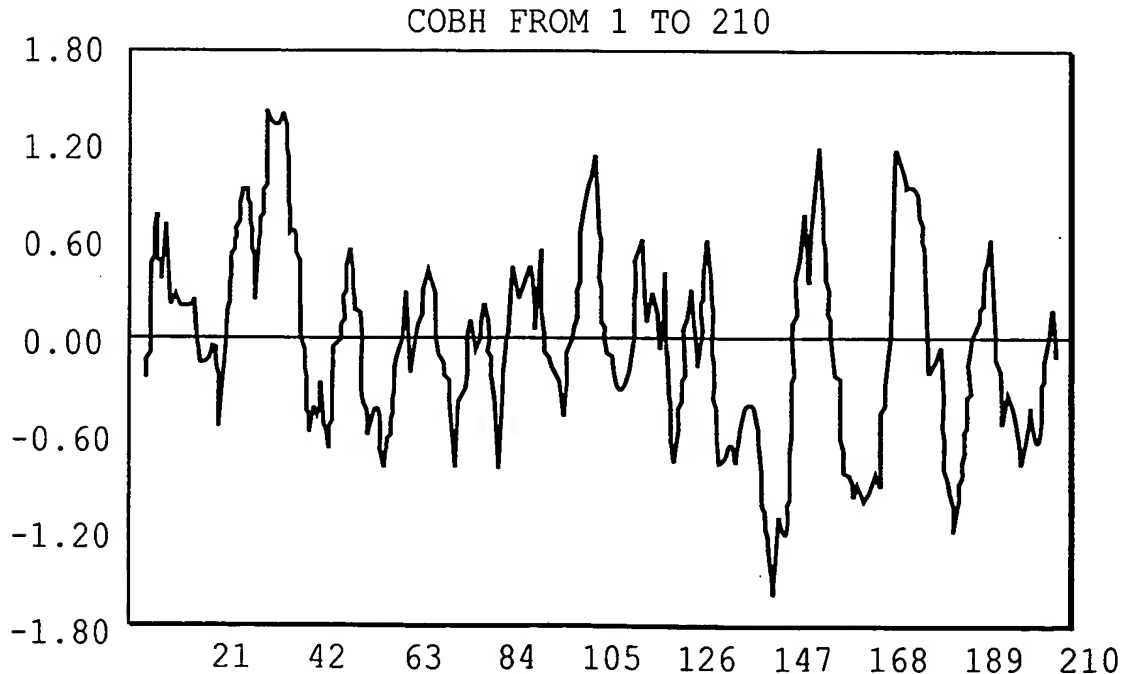


FIG. 16G

cobI GENE (SEQ ID NO: 19) AND COBI PROTEIN (SEQ ID NO: 20)
SEQUENCE OF THE 8753-BP FRAGMENT FROM 3631 TO 4368

MetSerGlyValGlyValGlyArgLeuIleGlyValGlyThrGlyProGlyAspProGluLeuLeuThrValLys
GTGAGCGGCGTCGGCGTGGGGCGCCTGATCGGTGTTGGGACCGCCCCGGTGATCCGGAACCTTTGACGGTCAAG
3631 3641 3651 3661 3671 3681 3691
AlaValLysAlaLeuGlyGlnAlaAspValLeuAlaTyrPheAlaLysAlaGlyArgSerGlyAsnGlyArgAla
GCGGTGAAGGCGCTCGGGCAAGCCGATGTGCTTGCCTATTTCCGCAAGGCCGGGCGAAGCGGTAAACGGCCGCGCG
3706 3716 3726 3736 3746 3756 3766
ValValGluGlyLeuLeuLysProAspLeuValGluLeuProLeuTyrTyrProValThrThrGluIleAspLys
GTGGTCGAGGGTCTGCTGAAGCCCGATCTTGTCGAGCTGCCGCTATACTATCCGGTGACGACCGAAATCGACAAG
3781 3791 3801 3811 3821 3831 3841
AspAspGlyAlaTyrLysThrGlnIleThrAspPheTyrAsnAlaSerAlaGluAlaValAlaAlaHisLeuAla
GACGATGGCGCCTACAAGACCCAGATACCGACTTCTACAATGCGTCGGCCGAAGCGGTAGCGGCGCATCTTGCC
3856 3866 3876 3886 3896 3906 3916
AlaGlyArgThrValAlaValLeuSerGluGlyAspProLeuPheTyrGlySerTyrMetHisLeuHisValArg
GCCGGGCGCACGGTCGCCGTGCTCAGTGAAGGCGACCCGCTGTTCTATGGTTCTACATGCATCTGCATGTGCGG
3931 3941 3951 3961 3971 3981 3991
LeuAlaAsnArgPheProValGluValIleProGlyIleThrAlaMetSerGlyCysTrpSerLeuAlaGlyLeu
CTCGCCAATCGTTTCCCGGTCGAGGTGATCCCCGGCATTACCGCCATGTCCGGCTGTTGGTCGCTTGCCGGCCTG
4006 4016 4026 4036 4046 4056 4066
ProLeuValGlnGlyAspAspValLeuSerValLeuProGlyThrMetAlaGluAlaGluLeuGlyArgArgLeu
CCGCTGGTGCAGGGCGACGACGTGCTCTCGGTGCTTCCGGGCACCATGGCCGAGGCCGAGCTCGGCCGAGGCTT
4081 4091 4101 4111 4121 4131 4141
AlaAspThrGluAlaAlaValIleMetLysValGlyArgAsnLeuProLysIleArgArgAlaLeuAlaAlaSer
GCGGATACCGAAGCCGCCGTGATCATGAAGGTCGGGCGCAATTTGCCGAAGATCCGTGGGCGCTCGCTGCCTCC
4156 4166 4176 4186 4196 4206 4216
GlyArgLeuAspGlnAlaValTyrValGluArgGlyThrMetLysAsnAlaAlaMetThrAlaLeuAlaGluLys
GGCCGTCTCGACCAGGCCGTCTATGTGAACGCGGCACGATGAAGAACGCGGCGATGACGGCTCTTGCGGAAAAG
4231 4241 4251 4261 4271 4281 4291
AlaAspAspGluAlaProTyrPheSerLeuValLeuValProGlyTrpLysAspArgPro***
GCCGACGACGAGGCGCCCTATTTCTCGCTGGTGCTCGTTCCCGGCTGGAAGGACCGACCATGA
4306 4316 4326 4336 4346 4356 4366

FIG. 16H

COBI PROTEIN			FIRST RESIDUE = 1		LAST RESIDUE = 245	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	5	2.04	735.35	2.84
2	LEU	L	28	11.43	3166.24	12.24
3	ILE	I	7	2.86	791.56	3.06
4	MET	M	7	2.86	917.28	3.54
5	VAL	V	25	10.20	2476.75	9.57
6	SER	S	10	4.08	870.30	3.36
7	PRO	P	14	5.71	1358.70	5.25
8	THR	T	12	4.90	1212.60	4.69
9	ALA	A	34	13.88	2415.36	9.33
10	TYR	Y	9	3.67	1467.54	5.67
11	*	*	0	0.00	0.00	0.00
12	HIS	H	3	1.22	411.18	1.59
13	GLN	Q	4	1.63	512.24	1.98
14	ASN	N	5	2.04	570.20	2.20
15	LYS	K	11	4.49	1408.99	5.44
16	ASP	D	15	6.12	1725.45	6.67
17	GLU	E	13	5.31	1677.52	6.48
18	CYS	C	1	0.41	103.01	0.40
19	TRP	W	2	0.82	372.16	1.44
20	ARG	R	14	5.71	2185.40	8.44
21	GLY	G	26	10.61	1482.52	5.73
22	-	-	0	0.00	0.00	0.00

RESIDUES = 245
 MOLECULAR WEIGHT = 25878.
 INDEX OF POLARITY (%) = 36.
 ISOELECTRIC POINT = 6.17
 OD 260 (1 mg/ml) = 0.512 OD 280 (1 mg/ml) = 0.843

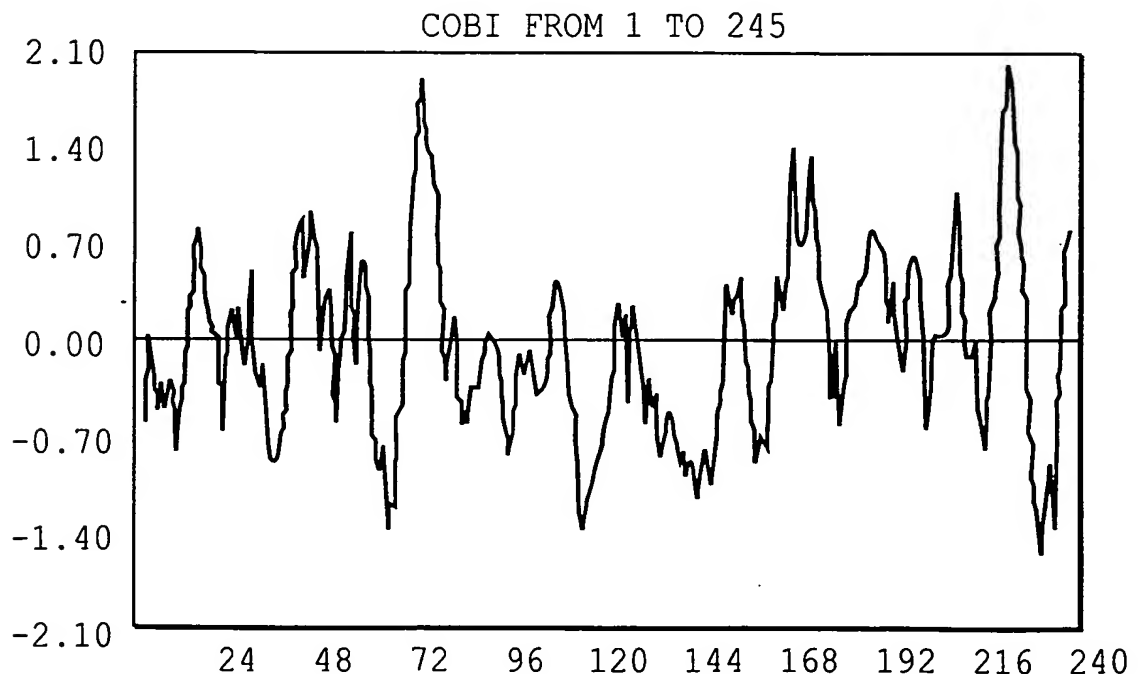


FIG. 16I

cobJ GENE (SEQ ID NO: 21) AND COBJ PROTEIN (SEQ ID NO: 22)
 SEQUENCE OF THE 8753-BP FRAGMENT FROM 4365 TO 5129

```

MetThrGlyThrLeuTyrValValGlyThrGlyProGlySerAlaLysGlnMetThrProGluThrAlaGluAla
ATGACCGGTACGCTCTATGTCGTCGGTACCGGACCGGGCAGCGCCAAGCAGATGACGCCGGAACGGCGGAAGCC
4365      4375      4385      4395      4405      4415      4425
ValAlaAlaAlaGlnGluPheTyrGlyTyrPheProTyrLeuAspArgLeuAsnLeuArgProAspGlnIleArg
GTTGCGGCCGCTCAGGAGTTTTACGGCTACTTCCCTATCTCGACCGGCTGAACCTCAGACCGGATCAGATCCGT
4440      4450      4460      4470      4480      4490      4500
ValAlaSerAspAsnArgGluGluLeuAspArgAlaGlnValAlaLeuThrArgAlaAlaAlaGlyValLysVal
GTCGCCTCGGACAACCGCGAGGAGCTCGATCGGGCACAGGTCGCGCTGACGCGGGCTGCGGCAGGCGTGAAGGTC
4515      4525      4535      4545      4555      4565      4575
CysMetValSerGlyGlyAspProGlyValPheAlaMetAlaAlaAlaValCysGluAlaIleAspLysGlyPro
TGCATGGTCTCCGGTGGCGATCCCGGTGTCTTTGCCATGGCGGCCCGCTCTGCGAGGCGATCGACAAGGGACCG
4590      4600      4610      4620      4630      4640      4650
AlaGluTrpLysSerValGluLeuValIleThrProGlyValThrAlaMetLeuAlaValAlaAlaArgIleGly
GCGGAATGGAAGTCGGTTGAACTGGTGATCACGCCCCGGCGTGACCGCGATGCTCGCCGTTGCCGCCCGCATCGGC
4665      4675      4685      4695      4705      4715      4725
AlaProLeuGlyHisAspPheCysAlaIleSerLeuSerAspAsnLeuLysProTrpGluValIleThrArgArg
GCGCCGCTCGGTCATGATTTCTGTGGGATCTCGCTTTCCGACAATCTGAAGCCCTGGGAAGTCATCACCCGGCGT
4740      4750      4760      4770      4780      4790      4800
LeuArgLeuAlaAlaGluAlaGlyPheValIleAlaLeuTyrAsnProIleSerLysAlaArgProTrpGlnLeu
CTCAGGCTGGCGGGCGGAAGCGGGCTTCGTCATTGCCCTCTACAATCCGATCAGCAAGGCGCGGCCCTGGCAGCTC
4815      4825      4835      4845      4855      4865      4875
GlyGluAlaPheGluLeuLeuArgSerValLeuProAlaSerValProValIlePheGlyArgAlaAlaGlyArg
GGTGAGGCCTTCGAGCTTCTGCGCAGCGTTCTGCCGGCAAGCGTTCGGGTCATCTTCGGCCGTGCGGCCGGGCGG
4890      4900      4910      4920      4930      4940      4950
ProAspGluArgIleAlaValMetProLeuGlyGluAlaAspAlaAsnArgAlaAspMetAlaThrCysValIle
CCGGACGAACGGATCGCGGTGATGCCGCTCGGCGAGGCCGATGCCAACC GCGCCGACATGGCGACCTGCGTCATC
4965      4975      4985      4995      5005      5015      5025
IleGlySerProGluThrArgIleValGluArgAspGlyGlnProAspLeuValTyrThrProArgPheTyrAla
ATCGGCTCGCCGAGACGCGCATCGTCGAGCGGACGGCCAACCCGATCTCGTCTACACACCGCGCTTCTATGCA
5040      5050      5060      5070      5080      5090      5100
GlyAlaSerGln***
GGGGCGAGCCAGTGA
5115      5125

```

FIG. 16J

COBJ PROTEIN			FIRST RESIDUE = 1		LAST RESIDUE = 254	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	8	3.15	1176.56	4.34
2	LEU	L	20	7.87	2261.60	8.35
3	ILE	I	13	5.12	1470.04	5.43
4	MET	M	7	2.76	917.28	3.39
5	VAL	V	23	9.06	2278.61	8.41
6	SER	S	11	4.33	957.33	3.53
7	PRO	P	18	7.09	1746.90	6.45
8	THR	T	12	4.72	1212.60	4.48
9	ALA	A	40	15.75	2841.60	10.49
10	TYR	Y	7	2.76	1141.42	4.21
11	*	*	0	0.00	0.00	0.00
12	HIS	H	1	0.39	137.06	0.51
13	GLN	Q	7	2.76	896.42	3.31
14	ASN	N	5	1.97	570.20	2.11
15	LYS	K	6	2.36	768.54	2.84
16	ASP	D	13	5.12	1495.39	5.52
17	GLU	E	16	6.30	2064.64	7.62
18	CYS	C	4	1.57	412.04	1.52
19	TRP	W	3	1.18	558.24	2.06
20	ARG	R	19	7.48	2965.90	10.95
21	GLY	G	21	8.27	1197.42	4.42
22	-	-	0	0.00	0.00	0.00

RESIDUES = 254
 MOLECULAR WEIGHT = 27088.
 INDEX OF POLARITY (%) = 35.
 ISOELECTRIC POINT = 5.43
 OD 260 (1 mg/ml) = 0.575 OD 260 (1 mg/ml) = 0.922

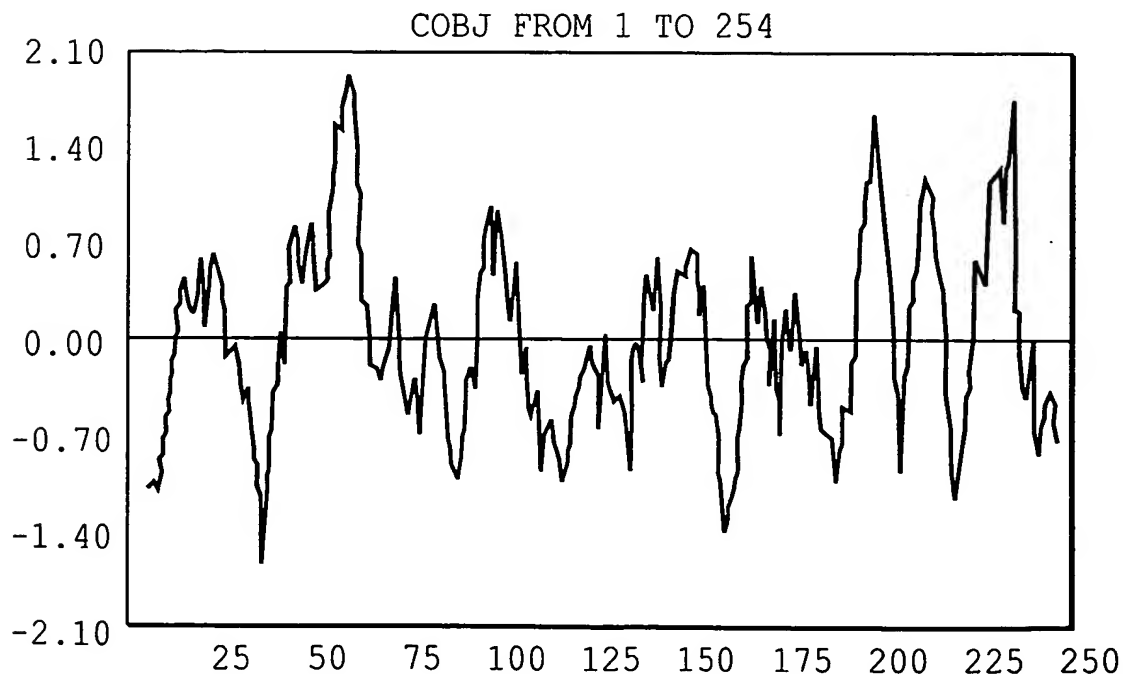


FIG. 16K

cobK GENE (SEQ ID NO: 23) AND COBK PROTEIN (SEQ ID NO: 24)
 SEQUENCE OF THE 8753-BP ECORI-ECORI FRAGMENT FROM 2861 TO 3646 ON
 THE COMPLEMENTARY STRAND

```

M A G S L F D T S A M E K P R I L I L G G T T E A
ATGGCGGGTTCGCTGTTGACACGTCAGCCATGGAAAAACCTCGTATTCTGATTCTGGGTGGCACCACCGAGGCA
2861      2871      2881      2891      2901      2911      2921      2931
R E L A R R L A E D V R Y D T A I S L A G R T A D
CGCGAACTCGCGCGCCGCTTGGCCGAAGATGTCCGCTACGACACCGCCATCTCGCTGGCCGGCCGACCGCGGAC
2936      2946      2956      2966      2976      2986      2996      3006
P R P Q P V K T R I G G F G G A D G L A H F V H D
CCGCGGCCGCGAGCCGGTCAAGACGCGCATCGGCGGCTTTGGCGGCGCCGATGGGCTGGCGCATTTCGTGCATGAC
3011      3021      3031      3041      3051      3061      3071      3081
E N I A L L V D A T H P F A A R I S H N A A D A A
GAAACATCGCGCTGCTGGTCGATGCGACGACCCCTTTGCCGACGCATTTTCGCACAACGCCGGGACGCGAGCG
3086      3096      3106      3116      3126      3136      3146      3156
Q R T G V A L I A L R R P E W V P L P G D R W T A
CAAAGAACCGGCGTTGCGCTTATCGCCCTCCGCCGACCGGAATGGGTGCCCCTGCCTGGCGACCGCTGGACTGCT
3161      3171      3181      3191      3201      3211      3221      3231
V D S V V E A V S A L G D R R R R V F L A I G R Q
GTCGATAGCGTTGTCGAGGCCGTCAGCGCGCTCGGTGATCGGCGACGCCGCGTCTTCCTGGCGATAGGTCGACAG
3236      3246      3256      3266      3276      3286      3296      3306
E A F H F E V A P Q H S Y V I R S V D P V T P P L
GAAGCTTTCCACTTCGAGGTCGCGCCGCGACAGCTACGTCATCCGCGAGCGTCGATCCGGTGACGCCGCCGCTT
3311      3321      3331      3341      3351      3361      3371      3381
N L P D Q E A I L A T G P F A E A D E A A L L R S
AATCTGCCCCGACCAGGAGGCGATCCTGGCGACCGGTCCCTTTGCGGAAGCCGACGAAGCCGCGTTGCTCAGGAGT
3386      3396      3406      3416      3426      3436      3446      3456
R Q I D V I V A K N S G G S A T Y G K I A A A R R
CGGCAGATCGATGTGATCGTCGCAAGAACAGCGGTGGCAGCGCCACCTACGGCAAGATTGCCGCGAGCGCGCCGG
3461      3471      3481      3491      3501      3511      3521      3531
L G I E V I M V E R R K P A D V P T V G S C D E A
CTCGGCATCGAGGTGATCATGGTCGAGCGGCGCAAGCCCGGGACGTGCCGAAGGTCGGCAGTTGCGACGAGGCA
3536      3546      3556      3566      3576      3586      3596      3606
L N R I A H W L A P A
CTCAACCGCATCGCTCACTGGCTCGCCCCTGCATGA
3611      3621      3631      3641

```

FIG. 16L

NAME = COBK FIRST RESIDUE = 1
 LAST RESIDUE = 261

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	8	3.07	1176.55	4.19
2	LEU	L	22	8.43	2487.85	8.86
3	ILE	I	16	6.13	1809.34	6.44
4	MET	M	3	1.15	393.12	1.40
5	VAL	V	21	8.05	2080.44	7.41
6	SER	S	12	4.60	1044.38	3.72
7	PRO	P	17	6.51	1649.90	5.88
8	THR	T	13	4.98	1313.62	4.68
9	ALA	A	42	16.09	2983.56	10.63
10	TYR	Y	3	1.15	489.19	1.74
11	*	*	0	0.00	0.00	0.00
12	HIS	H	7	2.68	959.41	3.42
13	GLN	Q	6	2.30	768.35	2.74
14	ASN	N	5	1.92	570.21	2.03
15	LYS	K	5	1.92	640.47	2.28
16	ASP	D	17	6.51	1955.46	6.96
17	GLU	E	15	5.75	1935.64	6.89
18	CYS	C	1	0.38	103.01	0.37
19	TRP	W	3	1.15	558.24	1.99
20	ARG	R	26	9.96	4058.63	14.45
21	GLY	G	19	7.28	1083.41	3.86
22	-	-	0	0.00	0.00	0.00

RESIDUES = 261
 MOLECULAR WEIGHT (MONOISOTOPIC) = 28078.7988
 MOLECULAR WEIGHT (AVERAGE) = 28096.0195
 INDEX OF POLARITY (%) = 40.64
 ISOELECTRIC POINT = 7.54
 OD 260 (1mg/ml) = 0.509 OD 280 (1mg/ml) = 0.721

HYDROPHILICITY PROFILE OF THE COBK PROTIEEN
 FROM 1 TO 261

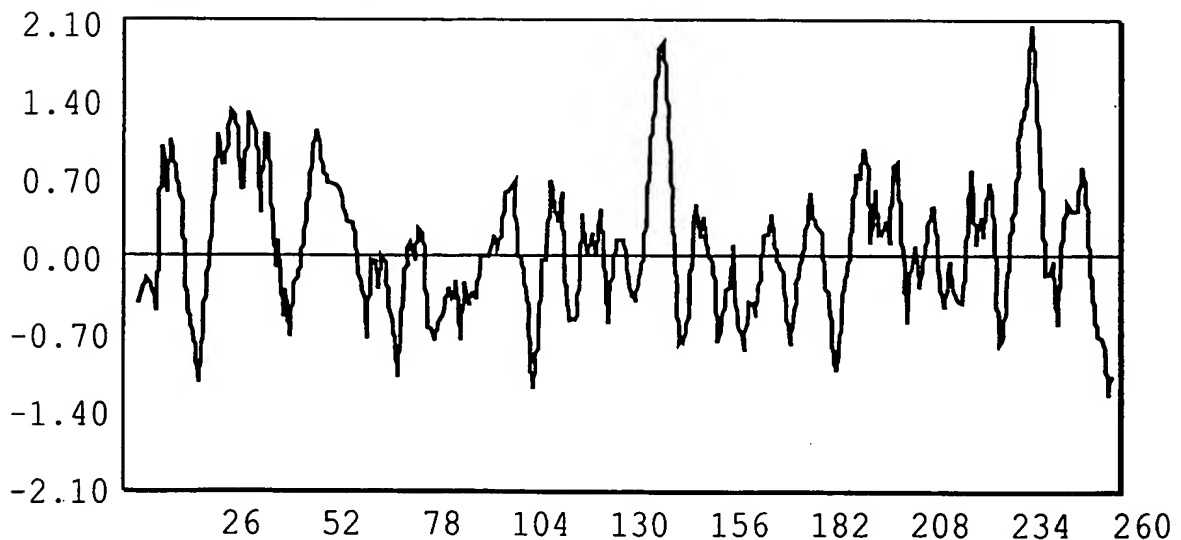


FIG. 16M

cobL GENE (SEQ ID NO: 25) AND COBL PROTEIN (SEQ ID NO: 26)
 SEQUENCE OF THE 8753-BP FRAGMENT FROM 5862 TO 7103

```

MetAlaAspValSerAsnSerGluProAlaIleValSerProTrpLeuThrValIleGlyIleGlyGluAspGly
ATGGCTGACGTGTCTGAACAGCGAACCCGCCATAGTCTCCCCCTGGCTGACCGTCATCGGTATCGGTGAGGATGGT
5862      5872      5882      5892      5902      5912      5922
ValAlaGlyLeuGlyAspGluAlaLysArgLeuIleAlaGluAlaProValValTyrGlyGlyHisArgHisLeu
GTAGCGGGTCTCGGCGACGAGGCCAAGCGGCTGATCGCCGAAGCGCCGGTCGTCTACGGCGGCCATCGTCATCTG
5937      5947      5957      5967      5977      5987      5997
GluLeuAlaAlaSerLeuIleThrGlyGluAlaHisAsnTrpLeuSerProLeuGluArgSerValValGluIle
GAGCTCGCCGCCTCCCTCATCACCGCGAAGCGCACAAATTGGCTAAGCCCCCTCGAACGCTCGGTCTGTCGAGATC
6012      6022      6032      6042      6052      6062      6072
ValAlaArgArgGlySerProValValValLeuAlaSerGlyAspProPhePhePheGlyValGlyValThrLeu
GTCGCGCGTCGCGGCAGCCCGGTGGTGGTGCTTGCCCTCGGGCGACCCGTTCTTCTTCGGCGTCGGCGTGACGCTG
6087      6097      6107      6117      6127      6137      6147
AlaArgArgIleAlaSerAlaGluIleArgThrLeuProAlaProSerSerIleSerLeuAlaAlaSerArgLeu
GCGCGCCGCATCGCCTCGGCCGAAATACGCACGCTTCCGGCGCCGTCGTCGATCAGTCTTGCCGCCTCGCGCCTC
6162      6172      6182      6192      6202      6212      6222
GlyTrpAlaLeuGlnAspAlaThrLeuValSerValHisGlyArgProLeuAspLeuValArgProHisLeuHis
GGCTGGGCGCTGCAGGATGCGACGCTCGTCTCCGTACATGGGCGGCCGCTGGATCTGGTGCGACCGCATTTCAT
6237      6247      6257      6267      6277      6287      6297
ProGlyAlaArgValLeuThrLeuThrSerAspGlyAlaGlyProArgAspLeuAlaGluLeuLeuValSerSer
CCGGGGGCGCGTGTGCTTACGCTCACGTCGGACGGTGCGGGTCCGCGAGACCTTGCCGAGCTTCTGGTTTCAAGC
6312      6322      6332      6342      6352      6362      6372
GlyPheGlyGlnSerArgLeuThrValLeuGluAlaLeuGlyGlyAlaGlyGluArgValThrThrGlnIleAla
GGCTTCGGTCAGTCGCGACTGACCGTGCTCGAAGCGCTGGGCGGCCGCGCGAACGGGTGACGACGCAGATCGCC
6387      6397      6407      6417      6427      6437      6447
AlaArgPheMetLeuGlyLeuValHisProLeuAsnValCysAlaIleGluValAlaAlaAspGluGlyAlaArg
GCGCGCTTCATGCTCGGCCTCGTGCATCCTTTGAACGTCTGCGCCATTGAGGTGGCGGCCGACGAGGGCGCGCGC
6462      6472      6482      6492      6502      6512      6522
IleLeuProLeuAlaAlaGlyArgAspAspAlaLeuPheGluHisAspGlyGlnIleThrLysArgGluValArg
ATCCTGCCGCTTGCCGCCGCGCGACGATGCGCTGTTCGAACATGACGGGCAGATACCAAGCGCGAGGTGCGG
6537      6547      6557      6567      6577      6587      6597
AlaLeuThrLeuSerAlaLeuAlaProArgLysGlyGluLeuLeuTrpAspIleGlyGlyGlySerGlySerIle
GCGCTGACGCTGTGCGCACTCGCACCGCGCAAGGGCGAACTGCTATGGGACATCGGCGGCGGCTCCGGCTCGATC
6612      6622      6632      6642      6652      6662      6672
GlyIleGluTrpMetLeuAlaAspProThrMetGlnAlaIleThrIleGluValGluProGluArgAlaAlaArg
GGCATCGAATGGATGCTCGCCGATCCGACCATGCAGCGATCACCATCGAGGTTGAGCCGGAGCGGGCAGCGCGC
6687      6697      6707      6717      6727      6737      6747

```

FIG. 16N

IleGlyArgAsnAlaThrMetPheGlyValProGlyLeuThrValValGluGlyGluAlaProAlaAlaLeuAla
ATCGGCCGCAACGCGACGATGTTCCGGCTGCCCCGGGCTGACGGTTGTCGAAGGCGAGGCGCCGGCGGCTTGCC
6762 6772 6782 6792 6802 6812 6822
GlyLeuProGlnProAspAlaIlePheIleGlyGlyGlyGlySerGluAspGlyValMetGluAlaAlaIleGlu
GGCCTGCCACAACCGGACGCGATCTTCATCGGCGGCGGCGGCAGCGAAGACGGCGTCATGGAAGCAGCGATCGAG
6837 6847 6857 6867 6877 6887 6897
AlaLeuLysSerGlyGlyArgLeuValAlaAsnAlaValThrThrAspMetGluAlaValLeuLeuAspHisHis
GCGCTCAAGTCAGGCGGACGGCTGGTTGCCAACGCGGTGACGACGGACATGGAAGCGGTGCTGCTCGATCATCAC
6912 6922 6932 6942 6952 6962 6972
AlaArgLeuGlyGlySerLeuIleArgIleAspIleAlaArgAlaGlyProIleGlyGlyMetThrGlyTrpLys
GCGCGGCTCGGCGGTTTCGCTGATCCGCATCGATATCGCGCGTGCAGGACCCATCGGCGGCATGACCGGCTGGAAG
6987 6997 7007 7017 7027 7037 7047
ProAlaMetProValThrGlnTrpSerTrpThrLysGly***
CCGGCCATGCCGGTCACCCAATGGTCGTGGACGAAGGGCTAA
7062 7072 7082 7092 7102

FIG. 160

COBL PROTEIN

FIRST RESIDUE = 1
 LAST RESIDUE = 413

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	8	1.94	1176.56	2.74
2	LEU	L	47	11.38	5314.76	12.39
3	ILE	I	26	6.30	2940.08	6.85
4	MET	M	9	2.18	1179.36	2.75
5	VAL	V	34	8.23	3368.38	7.85
6	SER	S	25	6.05	2175.75	5.07
7	PRO	P	24	5.81	2329.20	5.43
8	THR	T	21	5.08	2122.05	4.95
9	ALA	A	56	13.56	3978.24	9.27
10	TYR	Y	1	0.24	163.06	0.38
11	*	*	0	0.00	0.00	0.00
12	HIS	H	10	2.42	1370.60	3.19
13	GLN	Q	7	1.69	896.42	2.09
14	ASN	N	5	1.21	570.20	1.33
15	LYS	K	6	1.45	768.54	1.79
16	ASP	D	19	4.60	2185.57	5.09
17	GLU	E	27	6.54	3484.08	8.12
18	CYS	C	1	0.24	103.01	0.24
19	TRP	W	8	1.94	1488.64	3.47
20	ARG	R	28	6.78	4370.80	10.19
21	GLY	G	51	12.35	2908.02	6.78
22	-	-	0	0.00	0.00	0.00

RESIDUES = 413
 MOLECULAR WEIGHT = 42911.
 INDEX OF POLARITY (%) = 36.
 ISOELECTRIC POINT = 5.70
 OD 260 (1mg/ml) = 0.754 OD 280 (1mg/ml) = 1.064

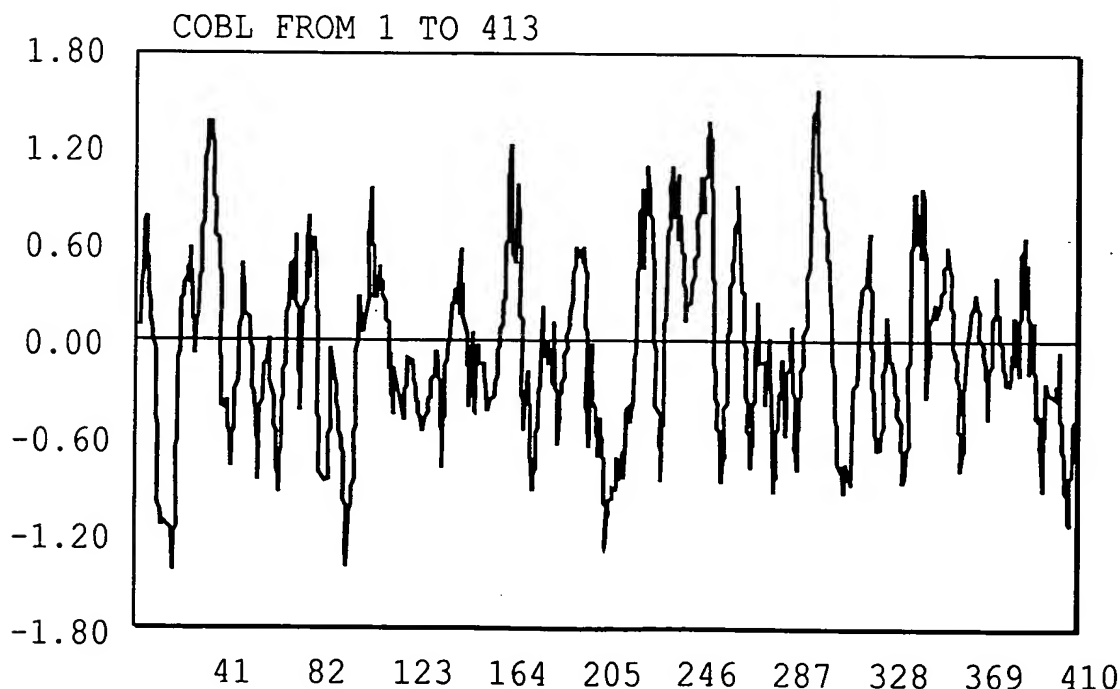


FIG. 16P

cobM GENE (SEQ ID NO: 27) AND COBM PROTEIN (SEQ ID NO: 28)
 SEQUENCE OF THE 8753-BP FRAGMENT FROM 7172 TO 7930

```

MetThrValHisPheIleGlyAlaGlyProGlyAlaAlaAspLeuIleThrValArgGlyArgAspLeuIleGly
ATGACGGTACATTTTCATCGGCGCCGGCCCGGGAGCCGCAGACCTGATCACGGTGCCTGGTCGCGACCTGATCGGG
7172      7182      7192      7202      7212      7222      7232
ArgCysProValCysLeuTyrAlaGlySerIleValSerProGluLeuLeuArgTyrCysProProGlyAlaArg
CGCTGCCCCGGTCTGCCTTTACGCCGGCTCGATCGTCTCGCCGGAGCTGCTGCGATATTGCCCGCCGGGCGCCCCG
7247      7257      7267      7277      7287      7297      7307
IleValAspThrAlaProMetSerLeuAspGluIleGluAlaGluTyrValLysAlaGluAlaGluGlyLeuAsp
ATTGTCGATACGGCGCCGATGTCCCTCGACGAGATCGAGGCGGAGTATGTGAAGCCGAAGCCGAAGGGCTCGAC
7322      7332      7342      7352      7362      7372      7382
ValAlaArgLeuHisSerGlyAspLeuSerValTrpSerAlaValAlaGluGlnIleArgArgLeuGluLysHis
GTGGCGCGGCTTCATTCGGGCGACCTTTCGGTCTGGAGTGTGTGGCCGAACAGATCCGCCGGCTCGAGAAGCAT
7397      7407      7417      7427      7437      7447      7457
GlyIleAlaTyrThrMetThrProGlyValProSerPheAlaAlaAlaAlaSerAlaLeuGlyArgGluLeuThr
GGCATCGCTATACGATGACGCCGGGCGTTCCTTCTTTCGGGCGGCGGCTTCAGCGCTCGGTGCGAATTGACC
7472      7482      7492      7502      7512      7522      7532
IleProAlaValAlaGlnSerLeuValLeuThrArgValSerGlyArgAlaSerProMetProAsnSerGluThr
ATTCGGGCGGTGGCCCAGAGCCTGGTGTGACCCGCTTTCGGGCCGCGCCTCGCCGATGCCGAACACAGAAACG
7547      7557      7567      7577      7587      7597      7607
LeuSerAlaPheGlyAlaThrGlySerThrLeuAlaIleHisLeuAlaIleHisAlaLeuGlnGlnValValGlu
CTTTCGGCTTTCGGCGCTACGGGATCGACGCTGGCAATCCACCTTGCGATCCATGCGCTTCAGCAGGTGGTCGAG
7622      7632      7642      7652      7662      7672      7682
GluLeuThrProLeuTyrGlyAlaAspCysProValAlaIleValValLysAlaSerTrpProAspGluArgVal
GAACTGACGCCGCTCTACGGTGCCGACTGCCCGGTGCGCCATCGTCGTCAGGCCTCTGGCCGGACGAACGCGTG
7697      7707      7717      7727      7737      7747      7757
ValArgGlyThrLeuGlyAspIleAlaAlaLysValAlaGluGluProIleGluArgThrAlaLeuIlePheVal
GTGCGCGGCACGCTCGGTGACATCGCCGCCAAGGTGGCGGAAGAGCCGATCGAGCGCACGGCGCTGATCTTCGTC
7772      7782      7792      7802      7812      7822      7832
GlyProGlyLeuGluAlaSerAspPheArgGluSerSerLeuTyrAspProAlaTyrGlnArgArgPheArgGly
GGTCCGGGGCTCGAAGCCTCCGATTTCCGTGAAAGCTCGCTCTACGATCCCGCCTATCAGCGGCGCTTCAGAGGG
7847      7857      7867      7877      7887      7897      7907
ArgGlyGlu
CGCGGCGAA
7922      7932      7942      7952      7962      7972      7982

```

FIG. 16Q

COBM PROTEIN			FIRST RESIDUE LAST RESIDUE		= 1 = 253	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	6	2.37	882.42	3.29
2	LEU	L	24	9.49	2713.92	10.11
3	ILE	I	15	5.93	1696.20	6.32
4	MET	M	4	1.58	524.16	1.95
5	VAL	V	22	8.70	2179.54	8.12
6	SER	S	18	7.11	1566.54	5.84
7	PRO	P	17	6.72	1649.85	6.15
8	THR	T	13	5.14	1313.65	4.89
9	ALA	A	35	13.83	2486.40	9.26
10	TYR	Y	7	2.77	1141.42	4.25
11	*	*	0	0.00	0.00	0.00
12	HIS	H	5	1.98	685.30	2.55
13	GLN	Q	5	1.98	640.30	2.39
14	ASN	N	1	0.40	114.04	0.42
15	LYS	K	4	1.58	512.36	1.91
16	ASP	D	11	4.35	1265.33	4.71
17	GLU	E	19	7.51	2451.76	9.13
18	CYS	C	4	1.58	412.04	1.53
19	TRP	W	2	0.79	372.16	1.39
20	ARG	R	19	7.51	2965.90	11.05
21	GLY	G	22	8.70	1254.44	4.67
22	-	-	0	0.00	0.00	0.00

RESIDUES = 253
 MOLECULAR WEIGHT = 26846.
 INDEX OF POLARITY (%) = 38.
 ISOELECTRIC POINT = 5.58
 OD 260 (1mg/ml) = 0.461 OD 280 (1mg/ml) = 0.724

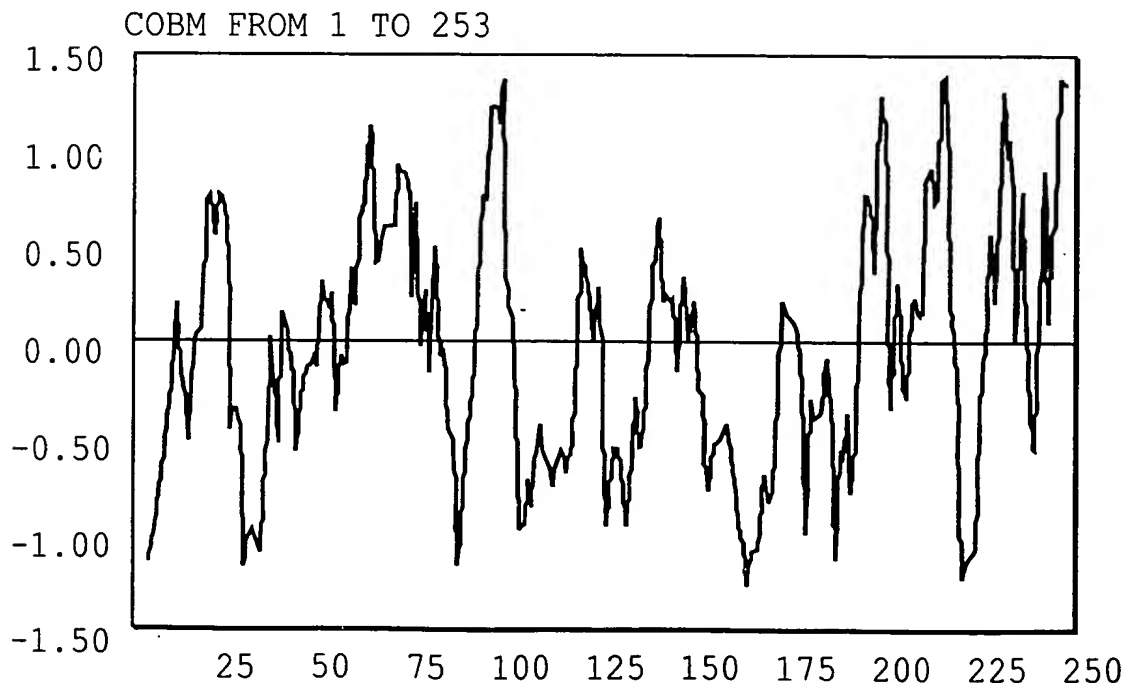
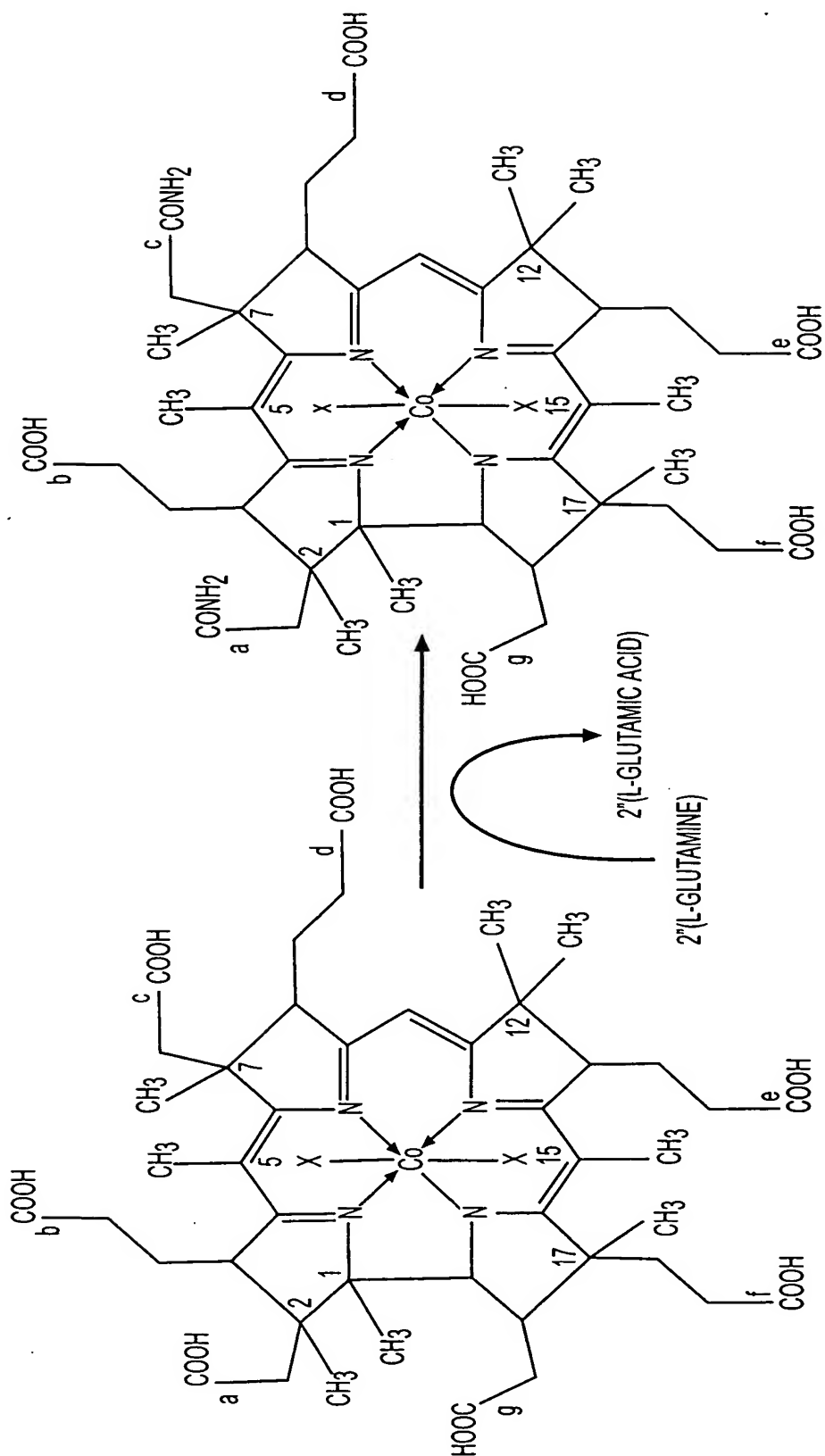


FIG. 16R



COBYRIC ACID DIAMIDE

FIG. 17

COBYRIC ACID

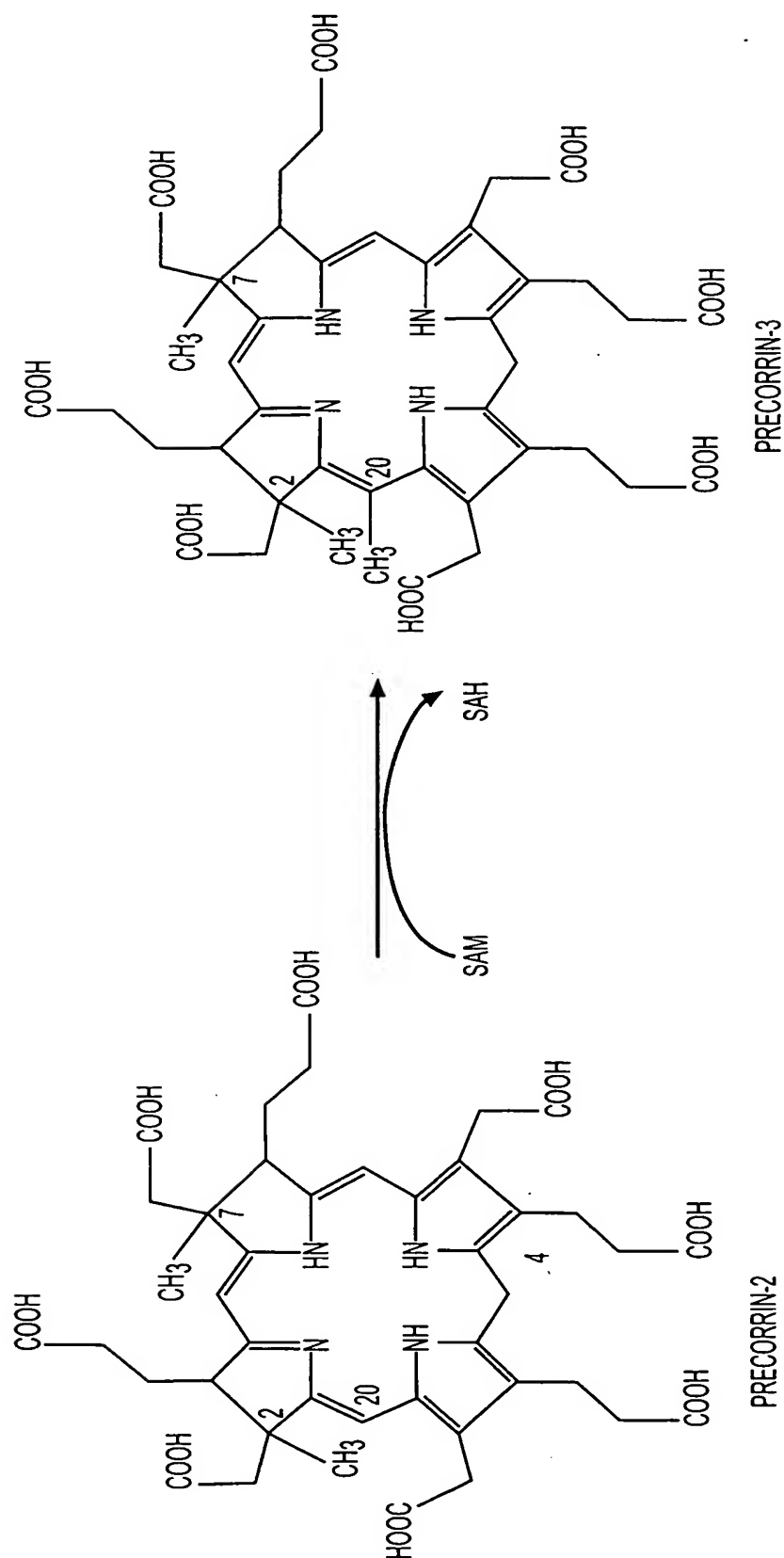
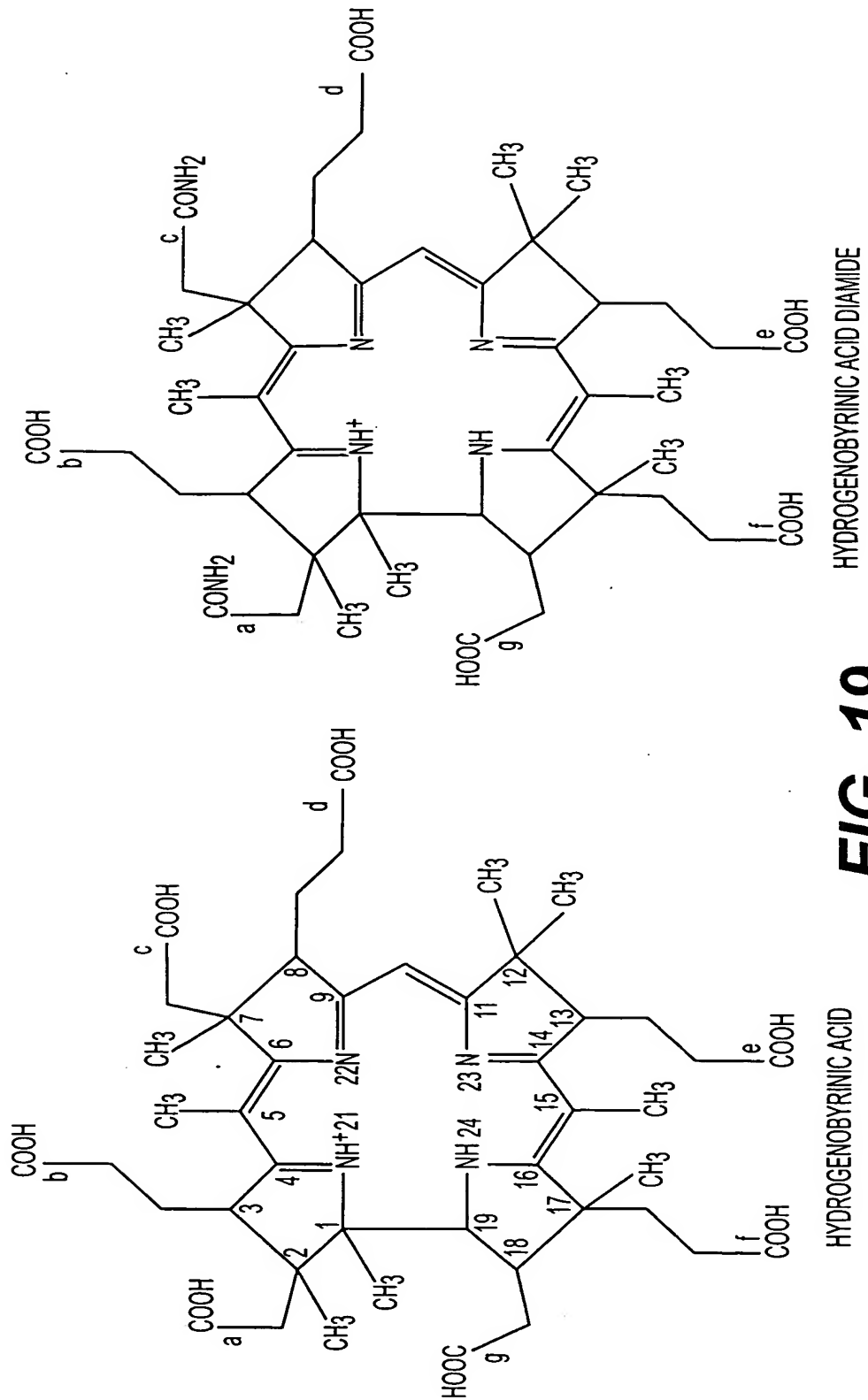


FIG. 18



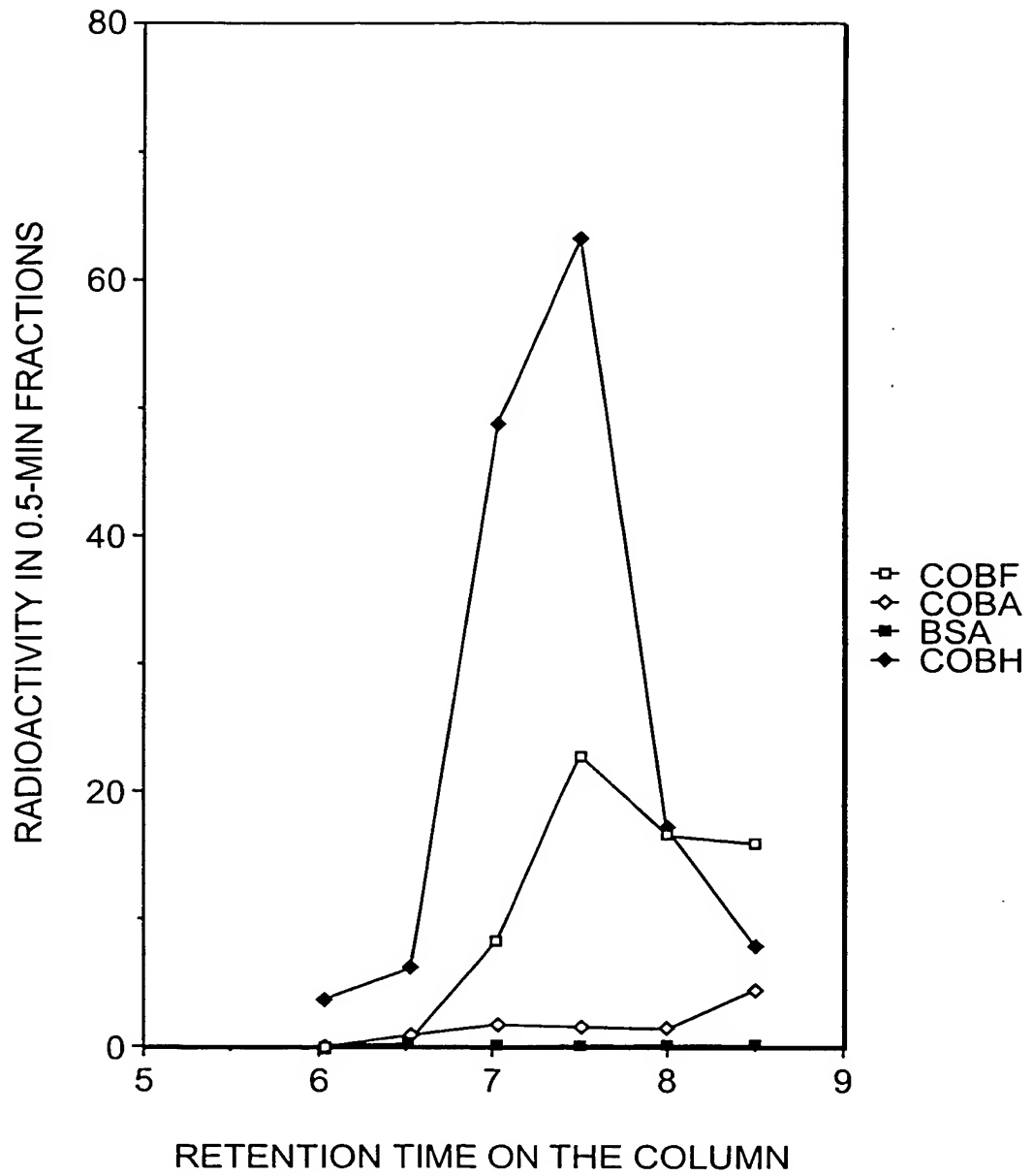


FIG. 20

```

10      20      30      40
COBA  LPALEKGSVWL-VGAGPGDPGLLTTHAANALRQADVIVH -(AMINO ACIDS 10-46 AND 88-157 OF SEQ ID NO:4)
- -- = - - - - - ===== --- = -- =====
COBI  VSGVGVGRL-IGVGTGPGDPPELLTVKAVKALGQADVLAY -(AMINO ACIDS 6-44 AND 99-164 OF SEQ ID NO:20)
10      20      30      40

```

REGION 1

```

90      100      110      120      130      140      150
COBA  LARAGNRVLRKGGDPVFVGRGGEALTLVEHQVPRIVPGITAGIGGLAYAGIPVTHREVNHAVTFLTG
== == = = == == -- -- - - - - = - - - - - - - - - - - - - - -
COBI  LA-AGRTVAVLSEGDPLFYGSYMHHLHVRLANR-FPVEVIPGITAMSGCWSLGLPLVQ-G-DDVLSVLPG
110      120      130      140      150      160

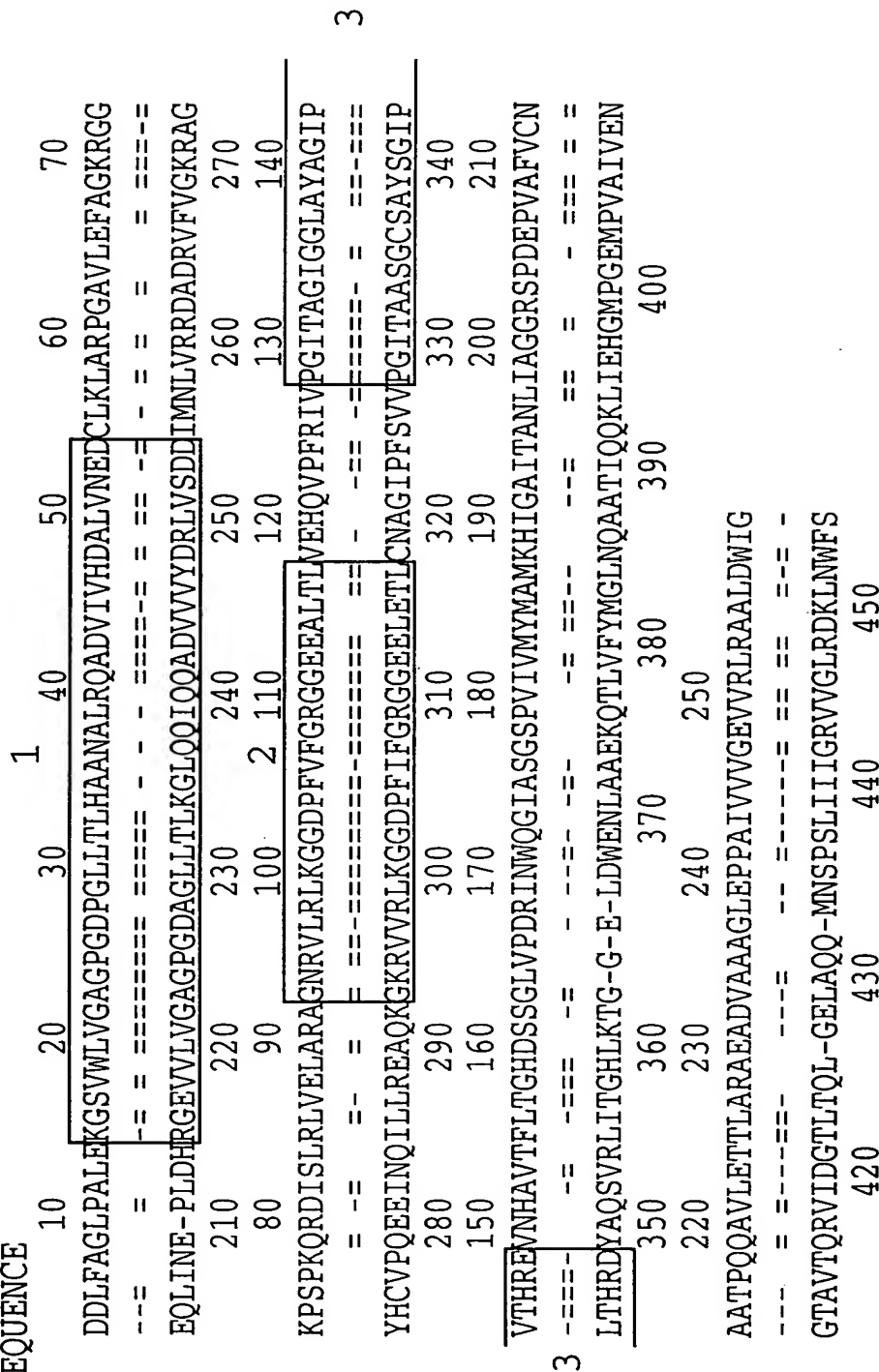
```

REGION 2

REGION 3

FIG. 21

ALIGNMENT BETWEEN THE PROTEINS COBA OF PSEUDOMONAS DENITRIFICANS
 AND CYSG OF ESCHERICHIA COLI.
 COBA SEQUENCE
 CYSG SEQUENCE



COBA SEQUENCE FROM 3 TO 259 - (AMINO ACIDS 3-259 OF SEQ ID NO:4)
 CYSG SEQUENCE FROM 204 TO 460 - (SEQ ID NO:55)
 STRICT HOMOLOGY OF AMINO ACIDS; 41.6%

FIG. 22

REGION 1

CYSG (209-248) EPLDHRGEVVLVGAGPGDAGLLTLKGLQQIQQADVVVYDR
 COBA (9-48) LPAL[E]KGSVWLVGAGPGDP[G]LLTLHAANALRQ[ADVI]VHDA
 COBI (1-40) VSGVGVRLLIGVGTGPGDPELLTLTKAVKALGQADVL[A]YFA
 COBF (1-40) MAEAGMRKILLIIGISGNPEHMTVQAINALNCADVLFIPT
 COBJ (1-40) MTGTLVVGTGPGSAKQMTPE[TA]EAVAAQEFYGYFPYLD
 COBL (6-45) NSEPAIIVSPW[L]TVIGIGEDGVAGLGDEAKALIAEAP[V]VYG
 COBM (1-34) MTVHFIGAGPGAADLITVRGRD[L]GRCPVCLYAG

REGION 2

CYSG (288-310) EAQKGKRVVRLKGGDPFIFGRGG- (CYSG IS SEQ ID NO:55)
 COBA (88-110) LARAGNRVLRLLKGGDPFVFGRGG- (COBA IS SEQ ID NO:4)
 COBI (98-120) HLAAGRTVAVLSEGGDPLFYGSYM- (COBI IS SEQ ID NO:20)
 COBF (60-83) VTRKDSRTVEFAVPVRRTEGVSY- (COBF IS SEQ ID NO:14)
 COBJ (68-80) RAAGVXVCMVSGGDPGVFAMAA- (COBJ IS SEQ ID NO:22)
 COBL (76-98) VARRGSPVVVVLASGDPFFGVGV- (COBL IS SEQ ID NO:26)
 COBM (69-91) AEAEGLDVARLHSGDLSVSAVA- (COBM IS SEQ ID NO:28)

REGION 3

CYSG (325-342) VVPGITAAAGCSAYSGIP
 COBA (125-142) IVPGITAGIGGLAYAGIP
 COBI (134-151) VIPGITAMSGCWSLAGLP
 COBF (85-102) SVDDWHAIAGIYEALLS
 COBJ (110-127) IITPGVTAMLAVARIGAP
 COBL (154-171) RVLTLTSDGAGPRDLAEL
 COBM (102-119) IAYTMTPGVPSPFAAASA

(AMINO ACID NOS. LISTED
 CORRESPOND TO IDENTIFIED AMINO
 ACIDS OF SEQ ID NOS.)

FIG. 23

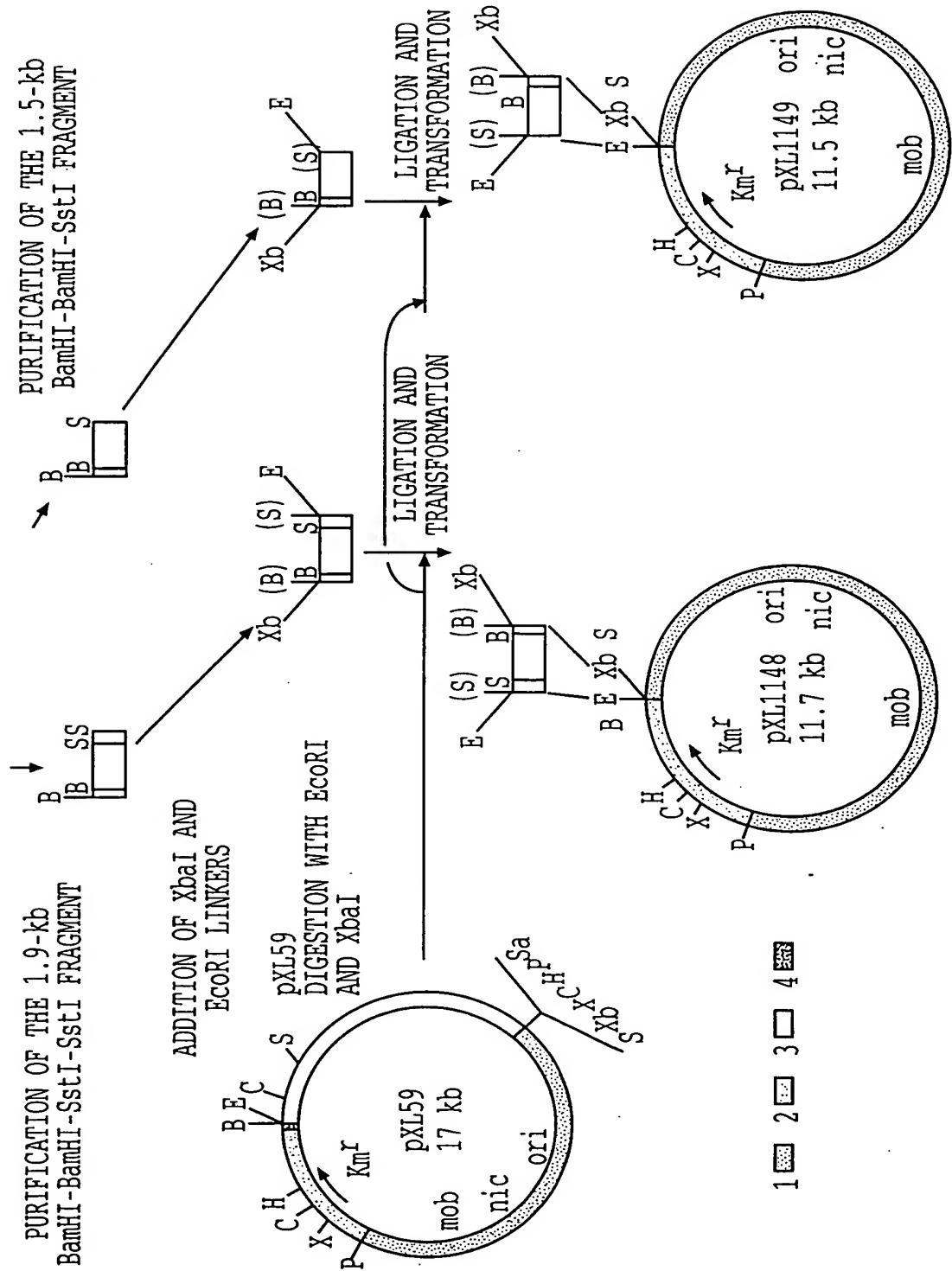


FIG. 24

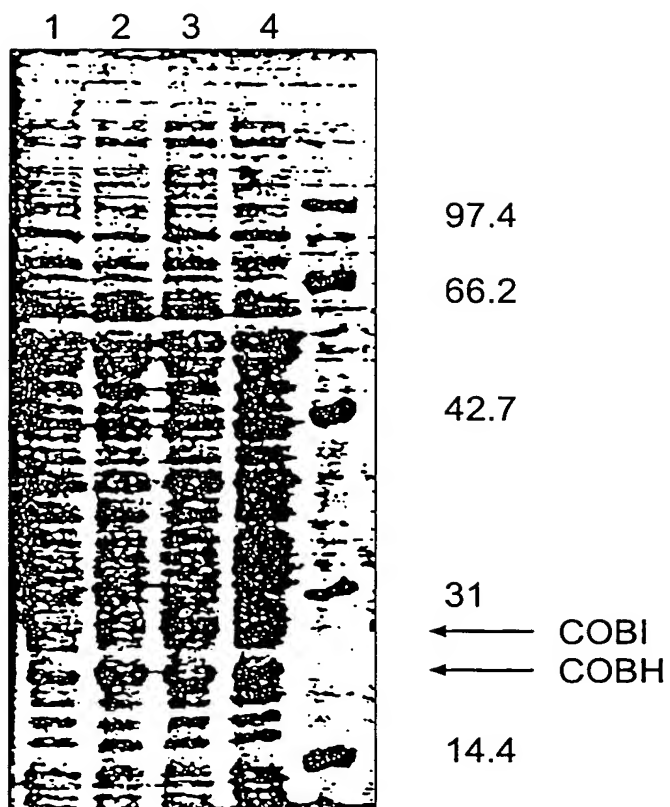


FIG. 25

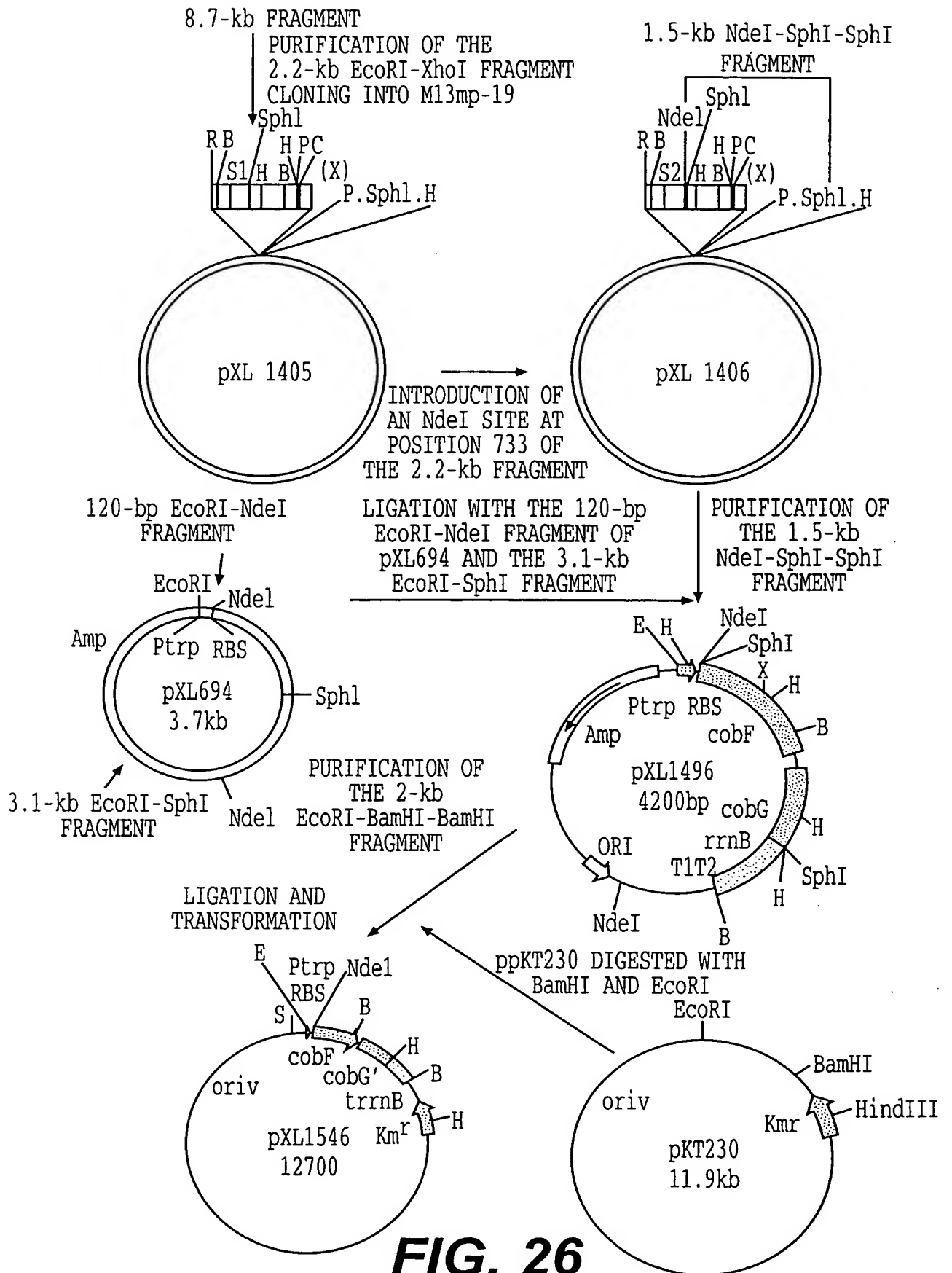


FIG. 26

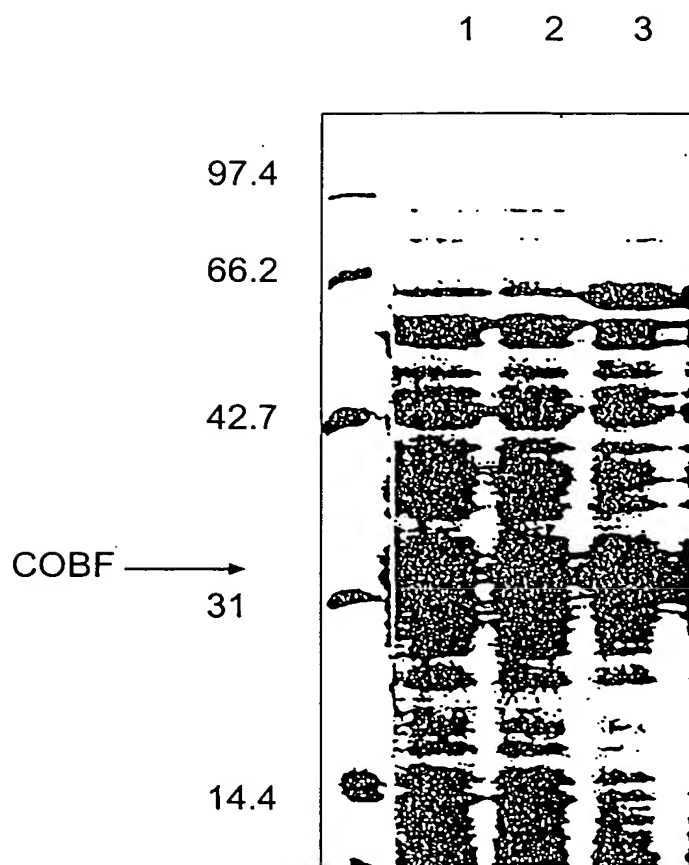


FIG. 27

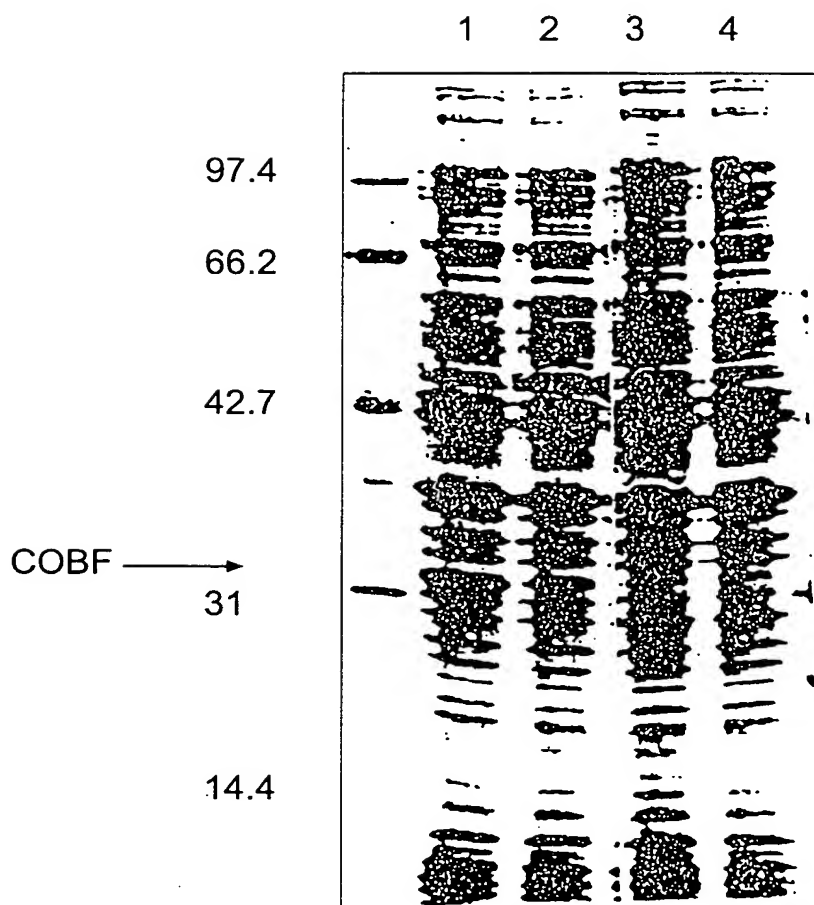


FIG. 28



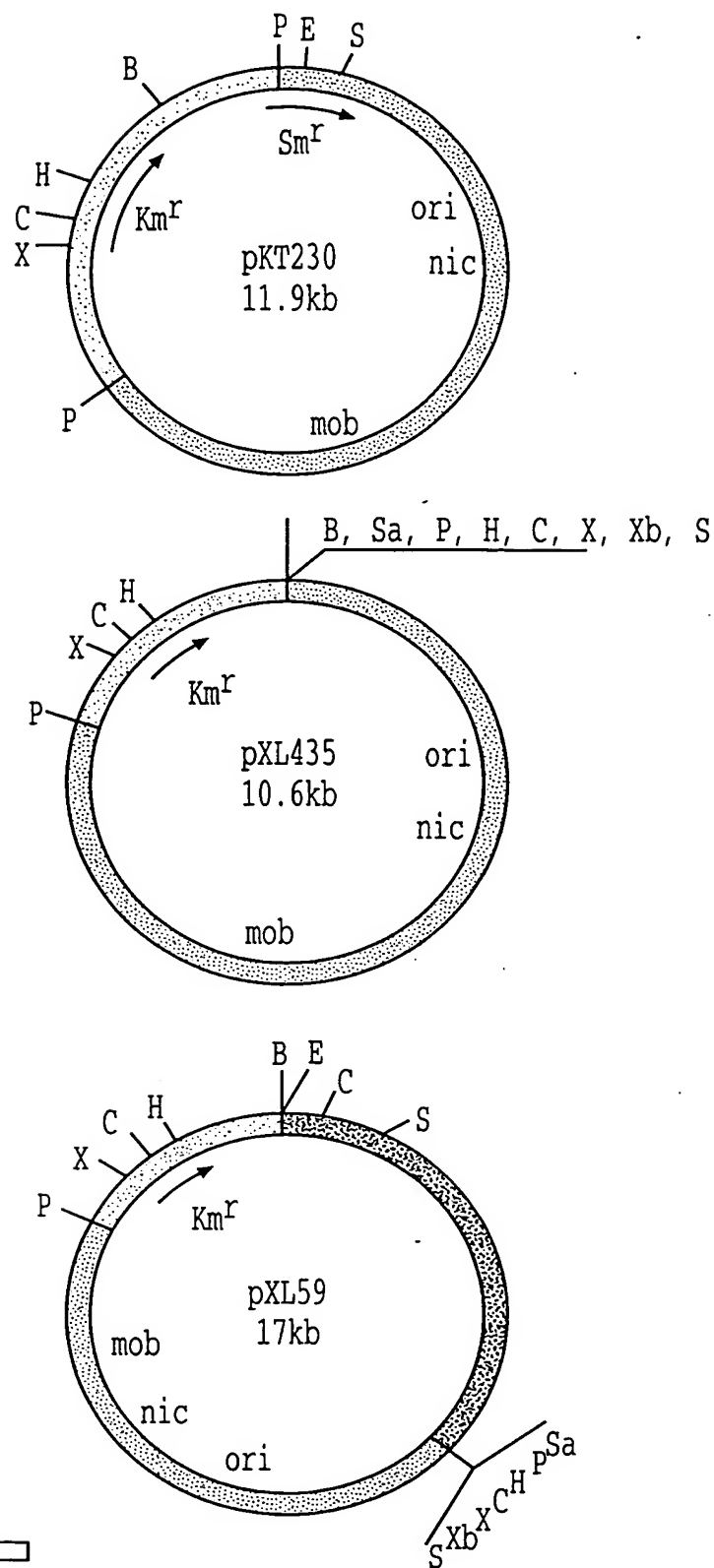
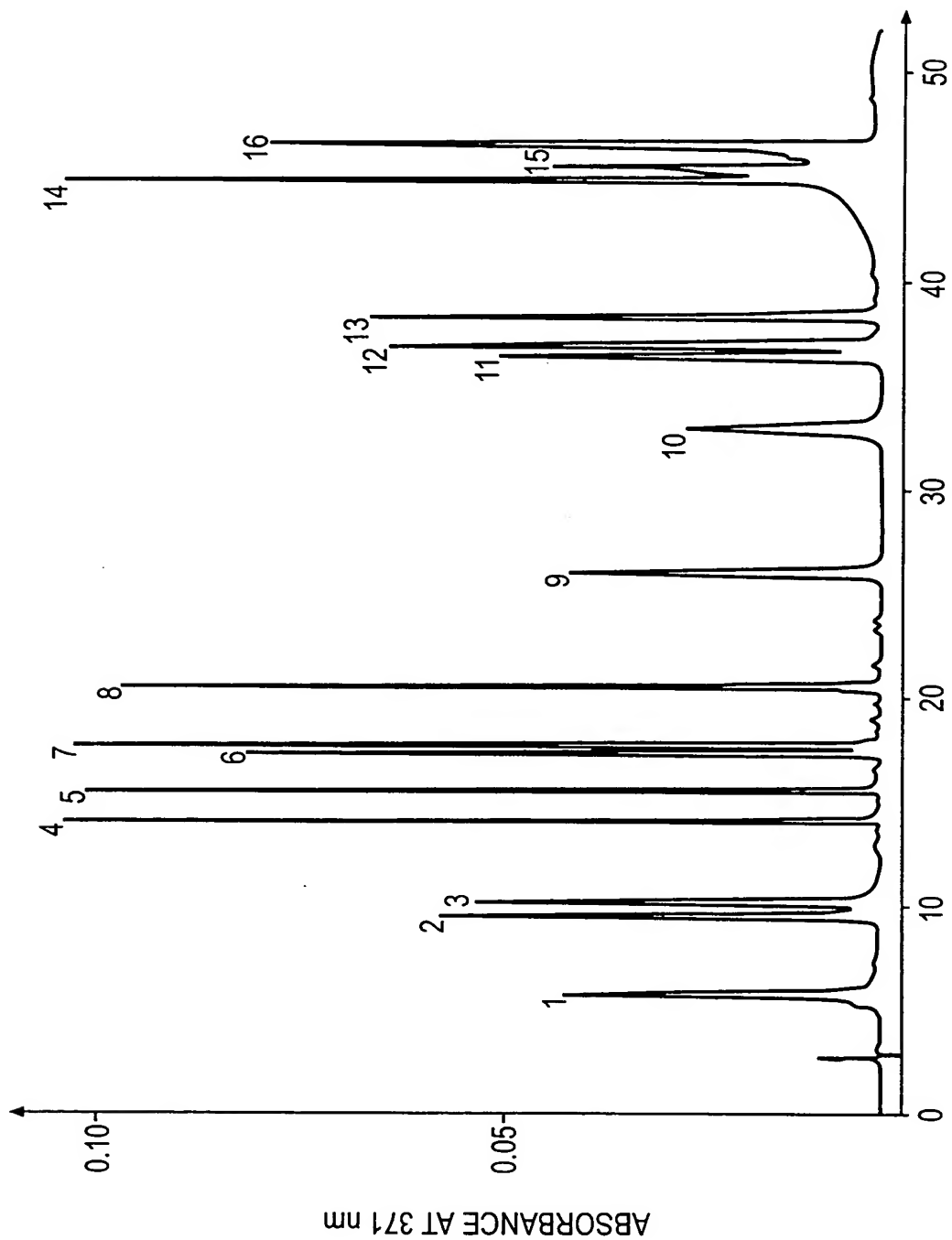


FIG. 30



TIME (MIN)
FIG. 31

10	20	30	40	50	60
GTCGACGAGT	ATGGTCAGGT	TCAGGGTCTG	GTGACGCTGG	AGGACATTCT	GGAGGAGATC
CAGCTGCTCA	TACCAGTCCA	AGTCCCAGAC	CACTGCGACC	TCCTGTAAGA	CCTCCTCTAG
70	80	90	100	110	120
GTCGGCGATA	TCGCCCATGA	GCACGACCTC	GACATTCAGG	GCGTGCGCCA	GGAAGCCGAT
CAGCCGCTAT	AGCGGCTACT	CGTGCTGGAG	CTGTAAGTCC	CGCACGCGGT	CCTTCGGCTA
130	140	150	160	170	180
GGCTCGATCG	TCGTCGATGG	CTCGGTGCCG	ATCCGCGATC	TCAACCGCGC	GCTCGACTGG
CCGAGCTAGC	AGCAGCTACC	GAGCCACGGC	TAGGCGCTAG	AGTTGGCGCG	CGAGCTGACC
190	200	210	220	230	240
TCGCTGCCGG	ACGAGGAGGC	GACGACGGTG	GCCGGTCTGG	TCATCCACGA	GTCCAAGAGC
AGCGACGGCC	TGCTCCTCCG	CTGCTGCCAC	CGGCCAGACC	AGTAGGTGCT	CAGGTTCTCG
250	260	270	280	290	300
ATTCCGGAGG	AGCGCCAGGC	CTTCACCTTC	CACGGCAAAC	GCTTCATCGT	GATGAAGCGG
TAAGGCCTCC	TCGCGGTCCG	GAAGTGAAG	GTGCCGTTTG	CGAAGTAGCA	CTACTTCGCC
310	320	330	340	350	360
GTGAAGAACC	GCATTACCAA	GCTGCGCATC	CGTCCGGCGG	AAGAGGGTGC	TCCGCCGGCG
CACTTCTTGG	CGTAATGGTT	CGACGCGTAG	GCAGGCCGCC	TTCTCCCACG	AGGCGGCCCG
370	380	390	400	410	420
TGATGGCCGC	GATTGCCTCT	ACCAGCGGGT	CGGCTCGCCG	GGGGCTGCCG	GCTCGACGGC
ACTACCGGCG	CTAACGGAGA	TGGTCGCCCCA	GCCGAGCGGC	CCCCGACGGC	CGAGCTGCCG
430	440	450	460	470	480
GAGCGCATGC	AGGCCGGCGT	CGAGTTCTGG	CTTCAGGAGA	TCATTGATGG	CGCGGTGGCG
CTCGCGTACG	TCCGGCCGCA	GCTCAAGACC	GAAGTCCTCT	AGTAACTACC	GCGCCACCGC
490	500	510	520	530	540
GGCGACACGG	CTCATGCCGG	CAAAGGCGCT	AGAAACGATA	CGCACCCGCA	TGTGGGACTC
CCGCTGTGCC	GAGTACGGCC	GTTTCCGCGA	TCTTTGCTAT	GCGTGGGCGT	ACACCCTGAG
550	560	570	580	590	600
GCCGGTACCA	TCGAAGCCCG	GCTGATGGCC	GGTATGCTGA	TGGCTCTCGT	TGATGACCTC
CGGCCATGGT	AGCTTCGGGC	CGACTACCGG	CCATACGACT	ACCGAGAGCA	ACTACTGGAG

FIG. 32A

610	620	630	640	650	660
GAGCCGTTCA	GGGTGGAAGG	CCTCGATCAG	CTTCTTTTCG	ATGGTCTCGG	TGAGCGACAT
CTCGGCAAGT	CCCACCTTCC	GGAGCTAGTC	GAAGAAAAGC	TACCAGAGCC	ACTCGCTGTA
670	680	690	700	710	720
TCTTCCGTCC	CATTTTGCTG	TTTGCTTGGC	GCCCCCTCGC	AGTTAAGAAC	CCGGTAATCG
AGAAGGCAGG	GTAAAACGAC	AAACGAACCG	CGGGGGAGCG	TCAATTCTTG	GGCCATTAGC
730	740	750	760	770	780
CTGGCACGGC	GGCGCAAAAT	GCCCGCACAA	AGCCAGCAAC	ATTCCGCTTT	GTCAATTCTT
GACCGTGCCG	CCGCGTTTTA	CGGGCGTGTT	TCGGTCGTTG	TAAGGCGAAA	CAGTTAAGAA
790	800	810	820	830	840
GTTGTGACTC	CCGCCAAACC	CCATAATGAG	CGCCATGAGA	CTCGATTCAA	AATACTTCGA
CAACACTGAG	GGCGGTTTGG	GGTATTACTC	GCGGTACTCT	GAGCTAAGTT	TTATGAAGCT
850	860	870	880	890	900
TCGCATTCTGA	ACCCGGCGCA	AGGTCGAGCC	GCAGCAGAGC	CGGCGGCTCC	CGTCTGTCAG
AGCGTAAGCT	TGGGCCGCGT	TCCAGCTCGG	CGTCGTCTCG	GCCGCCGAGG	GCAGACAGTC
910	920	930	940	950	960
TGGGACGGCT	GCGATCAGAA	GGGTGTGCAC	CGGGCGCCCCG	TCGGTCGCAA	CGCCGAGGGG
ACCCTGCCGA	CGCTAGTCTT	CCCACACGTG	GCCCGCGGGC	AGCCAGCGTT	GCGGCTCCCC
970	980	990	1000	1010	1020
CAGTACTTCA	TGTTCTGCTT	CGAGCACGTG	AAGGAATACA	ACAAGGGGCTA	CAACTTCTTC
GTCATGAAGT	ACAAGACGAA	GCTCGTGCAC	TTCTTATGT	TGTTCCCGAT	GTTGAAGAAG
1030	1040	1050	1060	1070	1080
TCCGGCCTCT	CCGACAGCGA	GGTCGCCCCG	TACCAGAAGG	AAGCGATCAC	CGGTCATCGG
AGGCCGGAGA	GGCTGTCGCT	CCAGCGGGCG	ATGGTCTTCC	TTCGCTAGTG	GCCAGTAGCC
1090	1100	1110	1120	1130	1140
CCCACCTGGA	CCGTCGGCGT	CAACAAGAAC	GCCAAGAACG	GCCCCGACCA	GTCGCAGACG
GGGTGGACCT	GGCAGCCGCA	GTTGTTCTTG	CGGTTCTTGC	CGGGCTGGGT	CAGCGTCTGC
1150	1160	1170	1180	1190	1200
CGGTCTGGCT	CTGCCGGCGC	CCAGGCCCGC	ATGCGCGATC	CCTTCGGCTT	TGTCAGCGAG
GCCAGACCGA	GACGGCCGCG	GGTCCGGGCG	TACGCGCTAG	GGAAGCCGAA	ACAGTCGCTC

FIG. 32B

1210	1220	1230	1240	1250	1260
GCGCGGGCGC	GATCCGGTCG	TCCCGAGCCA	CGCCAGCGCA	AGCTGAAGAC	GCTCGAGGCG
CGCGCCCGCG	CTAGGCCAGC	AGGGCTCGGT	GCGGTCGCGT	TCGACTTCTG	CGAGCTCCCG
1270	1280	1290	1300	1310	1320
AAGGCCTTCG	AAACGCTTGG	TCTCGGAGCC	TCGGCGACCA	CTGCCGACAT	CAAGGCGGCC
TTCCGGAAGC	TTTGCGAACC	AGAGCCTCGG	AGCCGCTGGT	GACGGCTGTA	GTTCCGCCGG
1330	1340	1350	1360	1370	1380
TACAAGGACC	TCGTCAAGAA	GCATCACCCC	GATGCCAATG	GCGGAGATAG	AGGATCGGAA
ATGTTCTCTG	AGCAGTTCTT	CGTAGTGGGG	CTACGGTTAC	CGCCTCTATC	TCCTAGCCTT
1390	1400	1410	1420	1430	1440
GAGCGTTTTTC	GCGCGGTTAT	TCAGGCCTAC	CAATTGTTAA	AACAGGCTGG	TTTCTGCTAA
CTCGCAAAAG	CGCGCCAATA	AGTCCGGATG	GTTAACAATT	TTGTCCGACC	AAAGACGATT
1450	1460	1470	1480	1490	1500
CAACCCGGAT	TAATACAGAA	GCACTTTTGC	AGGCGAATGC	GCGGGTGCCG	TCCGGTGGCC
GTTGGGCCTA	ATTATGTCTT	CGTGAAAACG	TCCGCTTACG	CGCCCACGGC	AGGCCACCGG
1510	1520	1530	1540	1550	1560
GCTCTGGAGA	CATGATGAGC	AAGATTGACC	TCGACATTTT	CAACCTCCCC	GACACCACGA
CGAGACCTCT	GTACTACTCG	TTCTAACTGG	AGCTGTAAAG	GTTGGAGGGG	CTGTGGTGCT
1570	1580	1590	1600	1610	1620
TTTCCGTCCG	GGAGGTTTTT	GGTATTGATA	CGGATTTGCG	CGTTCCTGCC	TATTCGAAGG
AAAGGCAGGC	CCTCCAAAAG	CCATAACTAT	GCCTAAACGC	GCAAGGACGG	ATAAGCTTCC
1630	1640	1650	1660	1670	1680
GCGACGCCTA	TGTCCCGGAT	CTGGATCCGG	ACTACCTCTT	CGACCGCGAA	ACGACGCTCG
CGCTGCGGAT	ACAGGGCCTA	GACCTAGGCC	TGATGGAGAA	GCTGGCGCTT	TGCTGCGAGC
1690	1700	1710	1720	1730	1740
CCATTCTCGC	AGGCTTCGCC	CACAACCGAC	GCGTGATGGT	GTCGGGCTAT	CACGGCACCG
GGTAAGAGCG	TCCGAAGCGG	GTGTTGGCTG	CGCACTACCA	CAGCCCGATA	GTGCCGTGGC
1750	1760	1770	1780	1790	1800
GCAAGTCCAC	CCATATCGAG	CAGGTCGCCG	CGCGCCTCAA	CTGGCCGTGC	GTGCGCGTCA
CGTTCAGGTG	GGTATAGCTC	GTCCAGCGGC	GCGCGGAGTT	GACCGGCACG	CACGCGCAGT

FIG. 32C

1810	1820	1830	1840	1850	1860
ACCTCGATAG	CCATGTCAGC	CGTATCGACC	TCGTCGGCAA	GGACGCGATC	GTCGTCAAGG
TGGAGCTATC	GGTACAGTCG	GCATAGCTGG	AGCAGCCGTT	CCTGCGCTAG	CAGCAGTTCC
1870	1880	1890	1900	1910	1920
ACGGCCTGCA	GGTCACCGAA	TTCAAGGACG	GCATCCTGCC	CTGGGCCTAC	CAGCACAATG
TGCCGACGCT	CCAGTGGCTT	AAGTTCCTGC	CGTAGGACGG	GACCCGGATG	GTCGTGTTAC
1930	1940	1950	1960	1970	1980
TCGCGCTCGT	CTTCGACGAA	TACGATGCCG	GCCGCCCGGA	CGTCATGTTC	GTCATCCAGC
AGCGCGAGCA	GAAGCTGCTT	ATGCTACGGC	CGGCGGGCCT	GCAGTACAAG	CAGTAGGTCG
1990	2000	2010	2020	2030	2040
GCGTGCTGGA	ATCCTCCGGC	CGCCTGACGC	TGCTCGACCA	GAGCCGTGTC	ATCCGTCCGC
CGCACGACCT	TAGGAGGCCG	GCGGACTGCG	ACGAGCTGGT	CTCGGCACAG	TAGGCAGGCG
2050	2060	2070	2080	2090	2100
ACCCGGCCTT	CCGCCTGTTT	GCGACCGCCA	ACACCGTCGG	CCTCGGCGAC	ACGACCGGCC
TGGGCCGGAA	GGCGGACAAA	CGCTGGCGGT	TGTGGCAGCC	GGAGCCGCTG	TGCTGGCCGG
2110	2120	2130	2140	2150	2160
TCTATCACGG	CACGCAGCAG	ATCAACCAGG	CGCAGATGGA	CCGCTGGTCC	ATCGTCACCA
AGATAGTGCC	GTGCGTCGTC	TAGTTGGTCC	GCGTCTACCT	GGCGACCAGC	TAGCAGTGGT
2170	2180	2190	2200	2210	2220
CGCTGAACTA	CCTGCCGCAC	GACAAGGAAG	TCGACATCGT	CGCCGCCAAG	GTCAAGGGCT
GCGACTTGAT	GGACGGCGTG	CTGTTCCTTC	AGCTGTAGCA	GCGGCGGTTC	CAGTTCCCGA
2230	2240	2250	2260	2270	2280
TCACCGCCGA	CAAGGGCCGC	GAGACCGTCT	CCAAGATGGT	ACGTGTCGCC	GACCTCACGC
AGTGGCGGCT	GTTCCCGGCG	CTCTGGCAGA	GGTTCTACCA	TGCACAGCGG	CTGGAGTGCG
2290	2300	2310	2320	2330	2340
GCGCAGCCTT	CATCAATGGC	GATCTCTCGA	CTGTCTATGAG	CCCGCGTACG	GTCATCACCT
CGCGTCGGAA	GTAGTTACCG	CTAGAGAGCT	GACAGTACTC	GGGCGCATGC	CAGTAGTGGA
2350	2360	2370	2380	2390	2400
GGGCCGAGAA	CGCCCACATC	TTCGGCGACA	TCGCTTTCGC	CTTCCGCGTG	ACCTTCCTCA
CCCGGCTCTT	GCGGGTGTAG	AAGCCGCTGT	AGCGAAAGCG	GAAGGCGCAC	TGGAAGGAGT

FIG. 32D

2410	2420	2430	2440	2450	2460
ACAAGTGC GA	CGAGCTGGAG	CGGGCGCTGG	TCGCCGAGCA	CTACCAGCGC	GCCTTCGGCA
TGTTACGCT	GCTCGACCTC	GCCCCGCGACC	AGCGGCTCGT	GATGGTCGCG	CGGAAGCCGT
2470	2480	2490	2500	2510	2520
TCGAGCTGAA	GGAATGCGCT	GCCAACATCG	TGCTCGAAGC	CACCGCCTGA	TCCCACGGCC
AGCTCGACTT	CCTTACGCGA	CGGTTGTAGC	ACGAGCTTCG	GTGGCGGACT	AGGGTGCCGG
2530	2540	2550	2560	2570	2580
TGCCGTCCCC	TTTGGGAGGG	CGGGTCATGA	CGCTGTGGCA	AACCGGATGA	CGCCCCACTG
ACGGCAGGGG	AAACCCTCCC	GCCCAGTACT	GCGACACCGT	TTGGCCTACT	GCGGGGTGAC
2590	2600	2610	2620	2630	2640
GGGCGCCGTC	GCCTCTGGCT	GAAGAAGGAA	CTGTCTGTGAG	CTCGAATTCG	AAGGCAAAGC
CCCGCGGCAG	CGGAGACCGA	CTTCTTCCTT	GACAGCACTC	GAGCTTAAGC	TTCCGTTTCG
2650	2660	2670	2680	2690	2700
CAACCACGCG	CGAGAATGCT	GCGGAACCGT	TCAAGCGGGC	GCTTTCCGGC	TGCATCCGAT
GTTGGTGCGC	GCTCTTACGA	CGCCTTGGCA	AGTTCGCCCC	CGAAAGGCCG	ACGTAGGCTA
2710	2720	2730	2740	2750	2760
CGATCGCGGG	CGATGCCGAG	GTGGAAGTCG	CCTTCGCCAA	CGAGCGGCCG	GGCATGACCG
GCTAGCGCCC	GCTACGGCTC	CACCTTCAGC	GGAAGCGGTT	GCTCGCCGGC	CCGTACTGGC
2770	2780	2790	2800	2810	2820
GCGAACGCAT	CCGTCTGCCG	GAACCTTTCCA	AGCGCCCGAC	CCTGCAGGAA	CTTGCCGTGA
CGCTTGCGTA	GGCAGACGGC	CTTGAAAGGT	TCGCGGGCTG	GGACGTCTTT	GAACGGCACT
2830	2840	2850	2860	2870	2880
CCCGCGGGCT	CGGTGACAGC	ATGGCGCTGC	GCAAGGCCTG	TACGCATGCG	CGGATCCAGC
GGGCGCCCCA	GCCACTGTCT	TACCGCGACG	CGTTCCGGAC	ATGCGTACGC	GCCTAGGTCT
2890	2900	2910	2920	2930	2940
GCACCATGTC	GCCGCAAGGG	GCGGACGCCC	GCGCGATCTT	CGATGCGGTG	GAGCAGGCTC
CGTGGTACAG	CGGCGTTCCC	CGCCTGCGGG	CGCGCTAGAA	GCTACGCCAC	CTCGTCCGAG
2950	2960	2970	2980	2990	3000
GTGTCGAGGC	GATCGGGTCG	TTGCGCATGG	CGGGTGTCGC	CAAGAACCTC	AACGTCATGC
CACAGCTCCG	CTAGCCCAGC	AACGCGTACC	GCCCCACAGC	GTTCTTGAG	TTGCAGTACG

FIG. 32E

3010	3020	3030	3040	3050	3060
TCGAAGAGAA	ATACGCCAAG	GCGAATTTTCG	CAACGATCGA	GCGCCAGGCG	GACGCGCCGC
AGCTTCTCTT	TATGCGGTTC	CGCTTAAAGC	GTTGCTAGCT	CGCGGTCCGC	CTGCGCGGCG
3070	3080	3090	3100	3110	3120
TCGGCGAGGC	CGTAGCGCTG	CTGGTGCGCG	AGAAGCTGAC	GGGCCAGAAG	CCGCCGGCGT
AGCCGCTCCG	GCATCGCGAC	GACCACGCGC	TCTTCGACTG	CCCGGTCTTC	GGCGGCCGCA
3130	3140	3150	3160	3170	3180
CTGCCGGCAA	GGTGCTCGAC	CTCTGGCGCG	AGTTCATCGA	GGGCAAGGCT	GCCGGCGACA
GACGGCCGTT	CCACGAGCTG	GAGACCGCGC	TCAAGTAGCT	CCCGTTCCGA	CGGCCGCTGT
3190	3200	3210	3220	3230	3240
TTGAGCACCT	GTCGTCGACG	ATCAACAACC	AGCAGGCCTT	TGCCCCGGGTC	GTTCGCGACA
AACTCGTGGA	CAGCAGCTGC	TAGTTGTTGG	TCGTCCGGAA	ACGGGCCCCAG	CAAGCGCTGT
3250	3260	3270	3280	3290	3300
TGCTGACCTC	GATGGAAGTC	GCCGAGAAAT	ACGGTGACGA	CGACAACGAG	CCGGACGAGC
ACGACTGGAG	CTACCTTCAG	CGGCTCTTTA	TGCCACTGCT	GCTGTTGCTC	GGCCTGCTCG
3310	3320	3330	3340	3350	3360
AGGAAAGCGA	GACCGACGAA	GACCAGCCGC	GCAGCCAGGA	GCAGGACGAG	AACGCCAGCG
TCCTTTCGCT	CTGGCTGCTT	CTGGTCGGCG	CGTCGGTCCT	CGTCCTGCTC	TTGCGGTCGC
3370	3380	3390	3400	3410	3420
ACGAGGAAGC	CGGCGACGAT	GCCGCACCCG	CCGACGAGAA	CCAGGCTGCC	GAAGAGCAGA
TGCTCCTTCG	GCCGCTGCTA	CGGCGTGGGC	GGCTGCTCTT	GGTCCGACGG	CTTCTCGTCT
3430	3440	3450	3460	3470	3480
TGGAAGAAGG	CGAGATGGAC	GGCGCGGAGA	TCTCCGACGA	CGATCTCCAG	GACGAAGGCG
ACCTTCTTCC	GCTCTACCTG	CCGCGCCTCT	AGAGGCTGCT	GCTAGAGGTC	CTGCTTCCGC
3490	3500	3510	3520	3530	3540
ACGAGGACAG	CGAAACGCCC	GGCGAGGTCA	AGCGTCCGAA	CCAGCCCTTC	GCCGACTTCA
TGCTCCTGTC	GCTTTGCGGG	CCGCTCCAGT	TCGCAGGCTT	GGTCGGGAAG	CGGCTGAAGT
3550	3560	3570	3580	3590	3600
ACGAGAAGGT	CGACTACGCC	GTCTTCACCC	GCGAGTTCGA	CGAGACGATT	GCCTCGGAAG
TGCTCTTCCA	GCTGATGCGG	CAGAAGTGGG	CGCTCAAGCT	GCTCTGCTAA	CGGAGCCTTC

FIG. 32F

3610	3620	3630	3640	3650	3660
AGCTTTGCGA	CGAGGCCGAG	CTCGACCGGC	TGCGCGCCTT	CCTCGACAAG	CAGCTTGCCC
TCGAAACGCT	GCTCCGGCTC	GAGCTGGCCG	ACGCGCGGAA	GGAGCTGTTC	GTCGAACGGG
3670	3680	3690	3700	3710	3720
ATCTTCAAGG	CGCGGTTCGGC	CGCCTTGCCA	ACCGGCTGCA	GCGCCGCCTG	ATGGCGCAGC
TAGAAGTTCC	GCGCCAGCCG	GCGGAACGGT	TGGCCGACGT	CGCGGCGGAC	TACCGCGTCG
3730	3740	3750	3760	3770	3780
AGAACCGCTC	CTGGGAGTTC	GATCTCGAAG	AGGGGTATCT	CGATTCGGCG	CGGCTTCAGC
TCTTGCGGAG	GACCCTCAAG	CTAGAGCTTC	TCCCCATAGA	GCTAAGCCGC	GCCGAAGTCG
3790	3800	3810	3820	3830	3840
GCATCATCAT	CGATCCGATG	CAGCCGCTTT	CCTTCAAGCG	CGAAAAGGAC	ACCAACTTCC
CGTAGTAGTA	GCTAGGCTAC	GTCGGCGAAA	GGAAGTTCGC	GCTTTTCCTG	TGGTTGAAGG
3850	3860	3870	3880	3890	3900
GCGATACCGT	CGTGACGCTG	CTGATCGACA	ATTCCGGCTC	GATGCGCGGC	CGTCCGATCA
CGCTATGGCA	GCACTGCGAC	GACTAGCTGT	TAAGGCCGAG	CTACGCGCCG	GCAGGCTAGT
3910	3920	3930	3940	3950	3960
CGGTTGCCGC	CACCTGCGCC	GATATCCTGG	CGCGCACGCT	CGAGCGCTGC	GGCGTCAAGG
GCCAACGGCG	GTGGACGCGG	CTATAGGACC	GCGCGTGCGA	GCTCGCGACG	CCGCAGTTCC
3970	3980	3990	4000	4010	4020
TCGAGATCCT	CGGTTTTACC	ACCAAGGCGT	GGAAGGGTGG	GCAGTCACGC	GAGAAGTGGC
AGCTCTAGGA	GCCAAAATGG	TGGTTCCGCA	CCTTCCCACC	CGTCAGTGCG	CTCTTCACCG
4030	4040	4050	4060	4070	4080
TGGCCGGCGG	CAAGCCACAG	GCCCCGGGTC	GCCTCAACGA	CCTGCGACAC	ATCGTCTACA
ACCGGCCGCC	GTTCGGTGTC	CGGGGCCCAG	CGGAGTTGCT	GGACGCTGTG	TAGCAGATGT
4090	4100	4110	4120	4130	4140
AGTCTGCCGA	CGCTCCGTGG	CGCCGGGCAC	GACGCAATCT	CGGCCTGATG	ATGCGGGAAG
TCAGACGGCT	GCGAGGCACC	GCGGCCCGTG	CTGCGTTAGA	GCCGGACTAC	TACGCCCTTC
4150	4160	4170	4180	4190	4200
GCCTGCTCAA	GGAAAACATC	GACGGCGAGG	CGTTGATTTG	GGCGCATGAG	CGGCTGATGG
CGGACGAGTT	CCTTTTGTAG	CTGCCGCTCC	GCAACTAAAC	CCGCGTACTC	GCCGACTACC

FIG. 32G

4210	4220	4230	4240	4250	4260
CGCGGCGCGA	ACAGCGGCGC	ATCCTGATGA	TGATTTTCGGA	CGGCGCGCCG	GTCGACGACT
GCGCCGCGCT	TGTCGCCGCG	TAGGACTACT	ACTAAAGCCT	GCCGCGCGGC	CAGCTGCTGA
4270	4280	4290	4300	4310	4320
CGACGCTGTC	GGTCAATCCA	GGAACTATC	TGGAGCGTCA	CCTGCGCGCG	GTCATCGAGC
GCTGCGACAG	CCAGTTAGGT	CCTTTGATAG	ACCTCGCAGT	GGACGCGCGC	CAGTAGCTCG
4330	4340	4350	4360	4370	4380
AGATCGAAAC	GCGCTCGCCG	GTGGAAGTGC	TGGCGATCGG	TATCGGCCAC	GACGTGACGC
TCTAGCTTTG	CGCGAGCGGC	CACCTTGACG	ACCGCTAGCC	ATAGCCGGTG	CTGCACTGCG
4390	4400	4410	4420	4430	4440
GCTACTATCG	CCGTGCCGTC	ACCATCGTCG	ATGCCGATGA	GCTTGCCGGC	GCGATGACCG
CGATGATAGC	GGCACGGCAG	TGGTAGCAGC	TACGGCTACT	CGAACGGCCG	CGCTACTGGC
4450	4460	4470	4480	4490	4500
AACAGCTGGC	CGCACTCTTC	GAGGACGAAA	GCCAGCGCCG	CGGTTCTTCG	CGTCTTCGCC
TTGTCGACCG	GCGTGAGAAG	CTCCTGCTTT	CGGTCGCGGC	GCCAAGAAGC	GCAGAAGCGG
4510	4520	4530	4540	4550	4560
GCGCCGGGTG	ATGCTTCCCC	CTTGGGGGCG	GTGGAACATC	GCCTCCGAGC	TGCCAATCGG
CGCGGCCCAC	TACGAAGGGG	GAACCCCCGC	CACCTTGTAG	CGGAGGCTCG	ACGGTTAGCC
4570	4580	4590	4600	4610	4620
CACCTGCACG	CATCGCTGGC	GGCCGAAGTC	AATTTACGGA	CATAGTTTTA	CAGTCTACCA
GTGGACGTGC	GTAGCGACCG	CCGGCTTCAG	TTAAATGCCT	GTATCAAAAT	GTCAGATGGT
4630	4640	4650	4660	4670	4680
AGCTACCATG	CGTGCGGGC	TCACTTTGAG	CGCACGCCGC	GTCATTCCCG	ATGCCCCCTG
TCGATGGTAC	GCACCGCCCC	AGTGAAACTC	GCGTGCGGCG	CAGTAAGGGC	TACGGGGGAC
4690	4700	4710	4720	4730	4740
AAGGTAATTC	TCTTGATGCT	TGGCCGCGGT	CTCCTAGCCC	TTTTCTCCT	GGCTTCGGCC
TTCCATGAAG	AGAAGTACGA	ACCGGCGCCA	GAGGATCGGG	AAAAGGAGGA	CCGAAGCCGG
4750	4760	4770	4780	4790	4800
TGCCCCGC					
ACGGGCCG					

FIG. 32H

10	20	30	40	50	60
GAGCTCATAG	AGCAGTTCCT	CGATCGACTT	CAGCAGTCGC	ATGAAATCCA	TGCCGTGCTC
CTCGAGTATC	TCGTCAAGGA	GCTAGCTGAA	GTCGTCAGCG	TACTTTAGGT	ACGGCACGAG
70	80	90	100	110	120
CCCTTGCTTC	TATGCGTGGC	ACGACCGCGC	GCCGGGGCCG	ATGCCGGTCA	GTCGCGCAGA
GGGAACGAAG	ATACGCACCG	TGCTGGCGCG	CGGCCCCGGC	TACGGCCAGT	CAGCGCGTCT
130	140	150	160	170	180
CGCAGCTCGT	CGGTACGCAT	CTGCAGCATC	TCCAGCGTCG	ACAGGAAGCT	CATGCCGAGC
GCGTCGAGCA	GCCATGCGTA	GACGTCGTAG	AGGTCGCAGC	TGTCCTTCGA	GTACGGCTCG
190	200	210	220	230	240
AGGCTCTGAT	CGAGCTTGCC	CTTGGCTGCG	ACCGTTGCGC	CGATGTTGCG	GCGGGTGATC
TCCGAGACTA	GCTCGAACGG	GAACCGACGC	TGGCAACGCG	GCTACAACGC	CGCCCACTAG
250	260	270	280	290	300
GGGCCGATCG	AGATCTCCTG	AAGCATCACG	GGGGCTGCCT	GGGCCCCGCC	ATTGGCTGTC
CCCGGCTAGC	TCTAGAGGAC	TTCGTAGTGC	CCCCGACGGA	CCCGGGCCGG	TAACCGACAG
310	320	330	340	350	360
ATGACCGTGA	CGATAAAGTT	GAGGTTGGCC	GGGTCGAGGC	CGATCTTTTC	CGCATCTTCA
TACTGGCACT	GCTATTTCAA	CTCCAACCGG	CCCAGCTCCG	GCTAGAAAAG	GCGTAGAAGT
370	380	390	400	410	420
TAGGTGAGCG	CGATGTTGCT	GGCGCCGGTA	TCGACCAGCA	TGCTGATGTC	CTTGCCGTCG
ATCCACTCGC	GCTACAACGA	CCGCGGCCAT	AGCTGGTTCG	ACGACTACAG	GAACGGCAGC
430	440	450	460	470	480
ACCGTCGCAG	TGGTCTCGAA	ATGACCGTTC	AGCATCTTCT	GCAGCACCAC	TTCCTGCTGT
TGGCAGCGTC	ACCAGAGCTT	TACTGGCAAG	TCGTAGAAGA	CGTCGTGGTG	AAGGACGACA
490	500	510	520	530	540
CCCTCGCTGT	CAGTGATGAT	GGTGGCGCGG	CCGGGGATGA	GGCCGGCGAG	CAGGCGGTTA
GGGAGCGACA	GTCACTACTA	CCACCGCGCC	GGCCCCTACT	CCGGCCGCTC	GTCCGCCAAT
550	560	570	580	590	600
CCGAAGCCCT	CCAACTCGAA	GCGGTAGACA	TAGGCCGAGA	CCAGCGCCAG	AACGACGAAG
GGCTTCGGGA	GGTTGAGCTT	CGCCATCTGT	ATCCGGCTCT	GGTCGCGGTC	TTGCTGCTTC

FIG. 33A

610	620	630	640	650	660
AGCCAGATGG	CGATCTGACG	CAGGCCTTCG	CCGAAGCGGT	GGCGGCTCTG	CAGGATGCCG
TCGGTCTACC	GCTAGACTGC	GTCCGGAAGC	GGCTTCGCCA	CCGCCGAGAC	GTCCTACGGC
670	680	690	700	710	720
GCGCCGATCA	GCGTGGCGAT	GGCGCCGAGC	GAGACCAGTT	GCCCCAACTG	GTCATTGGCA
CGCGGCTAGT	CGCACCGCTA	CCGCGGCTCG	CTCTGGTCAA	CGGGCTTGAC	CAGTAACCGT
730	740	750	760	770	780
AGCCCCATGG	TGCGGCCGGT	GTCGTGGTTG	ATGATCAGCA	GGATGAGGCC	GATGGCCAGG
TCGGGGTACC	ACGCCGGCCA	CAGCACCAAC	TACTAGTCGT	CCTACTCCGG	CTACCGGTCC
790	800	810	820	830	840
ATCGAGAGCA	GGATGGCAAG	ACGGGTCATG	CTTCGCCGCG	TTCCCTCGCC	ATGCGCGTGC
TAGCTCTCGT	CCTACCGTTC	TGCCCAGTAC	GAAGCGGCGC	AAGGGAGCGG	TACGCGCACG
850	860	870	880	890	900
GTCGGGTTTC	GCGCCGCGGC	TTGCGTTCGA	CGGTCTCAAG	CCGTGCAGGC	AACGCGCTCA
CAGCCCAAAG	CGCGGCGCCG	AACGCAAGCT	GCCAGAGTTC	GGCACGTCCG	TTGCGCGAGT
910	920	930	940	950	960
TGATCGCGCG	GCGTTCGGCA	TCGGTATAGA	GCGTCCAGCG	TCCGACTTCG	TCGCGGGTAC
ACTAGCGCGC	CGCAAGCCGT	AGCCATATCT	CGCAGGTCGC	AGGCTGAAGC	AGCGCCCATG
970	980	990	1000	1010	1020
GGCCGCAGCC	GAAACAGTAG	CCGGTCTTGT	CATCGATCGA	ACAGACGAGA	ATGCAGGGAG
CCGGCGTCGG	CTTTGTCATC	GGCCAGAACA	GTAGCTAGCT	TGTCTGCTCT	TACGTCCCTC
1030	1040	1050	1060	1070	1080
ATTCCATGGG	CGTGCTCAGT	TTTCCCTTGA	TATATCGATG	TTTCAAACCG	TCAGCGCAAG
TAAGGTACCC	GCACGAGTCA	AAAGGGAAC	ATATAGCTAC	AAAGTTTGGC	AGTCGCGTTC
1090	1100	1110	1120	1130	1140
GGCACCGAGC	ACGGCGATTT	CGGTCAAGTTG	CTGCGTCGCC	CCGATCGTGT	CGCCCGTTTG
CCGTGGCTCG	TGCCGCTAAA	GCCAGTCAAC	GACGCAGCGG	GGCTAGCACA	GCGGGCAAAC
1150	1160	1170	1180	1190	1200
TCCGCCGATC	TTGCGCATCG	CCAGCCGAGC	GAAGCCCTTG	ACCGTGGCAA	GGAATGCGAC
AGGCGGCTAG	AACGCGTAGC	GGTCGGCTCG	CTTCGGGAAC	TGGCACCGTT	CCTTACGCTG

FIG. 33B

1210	1220	1230	1240	1250	1260
GAGCGCCGCG	ATGACGCCGA	GCGCCGGGAC	CTGCGCGAGA	TAGAAGAGCA	GCATTGCGAC
CTCGCGGCGC	TACTGCGGCT	CGCGGCCCTG	GACGCGCTCT	ATCTTCTCGT	CGTAACGCTG
1270	1280	1290	1300	1310	1320
AAGAAGTCCG	AAGGCAAGCG	CGAAGCGCGT	GGCCGCCGGT	TCCGGCTCGC	CAGCCGAGGC
TTCTTCAGGC	TTCCGTTCGC	GCTTCGCGCA	CCGGCGGCCA	AGGCCGAGCG	GTCGGCTCCG
1330	1340	1350	1360	1370	1380
CGCGACGCCG	CTGCTGCGCG	CCGGCGGAAG	CGACGACCAG	TGCCAGACCA	TGGCGGCGCG
GCGCTGCGGC	GACGACGCGC	GGCCGCCCTT	GCTGCTGGTC	ACGGTCTGGT	ACCGCCGCGC
1390	1400	1410	1420	1430	1440
GCTGAGGCAC	GCTGCGCCAA	GGATCGCCAT	GGCGGCGCCC	AGCGGCGAAA	AGAGCGGCAG
CGACTCCGTG	CGACGCGGTT	CCTAGCGGTA	CCGCCGCGGG	TCGCCGCTTT	TCTCGCCGTC
1450	1460	1470	1480	1490	1500
GATCGAGGCG	AACGCCGAGA	CGCGCAGGCC	GAAGGAGAGG	ATGAGGGCGA	CGGCCGCATA
CTAGCTCCGC	TTGCGGCTCT	GCGCGTCCGG	CTTCCTCTCC	TACTCCCGCT	GCCGGCGTAT
1510	1520	1530	1540	1550	1560
GGTGCCGATG	CGGCTGTCCT	TCATGATCGC	AAGCGCCGCT	TCGCGGTGCG	GACCGCCGCC
CCACGGCTAC	GCCGACAGGA	AGTACTAGCG	TTGCGGGCGA	AGCGCCAGCG	CTGGCGGCGG
1570	1580	1590	1600	1610	1620
AAAGCCATCG	GCCGTGTCGC	CAAGCCCGTC	TTCGTGCAGT	GCGCCGGTGA	CAAGCGCCTG
TTTCGGTAGC	CGGCACAGCG	GTTCGGGCAG	AAGCACGTCA	CGCGGCCACT	GTTTCGCGGAC
1630	1640	1650	1660	1670	1680
GATGGCGACG	ACGACAAAGG	CGGCAAAGAG	CGAGCTCACC	TGCAGCGCCA	TGAGGGCCAT
CTACCGCTGC	TGCTGTTTCC	GCCGTTTCTC	GCTCGAGTGG	ACGTCGCGGT	ACTCCCGGTA
1690	1700	1710	1720	1730	1740
GGCGACGGCC	GCCGATGGCA	GTGCGATCGC	CAGGCCGGCG	AACGGGAAGG	CGCGCACGGC
CCGCTGCCGG	CGGCTACCGT	CACGCTAGCG	GTCCGGCCGC	TTGCCCTTCC	GCGCGTGCCG
1750	1760	1770	1780	1790	1800
ACGGCTCAAG	CGCCCGTCAT	AACCTTCGAA	ATGACGCGCA	GGCATCGGGA	TGCGGCTGAG
TGCCGAGTTC	GCGGGCAGTA	TTGGAAGCTT	TACTGCGCGT	CCGTAGCCCT	ACGCCGACTC

FIG. 33C

1810	1820	1830	1840	1850	1860
AAAGCCGATC	GACCGCGCCA	CATCGTCACA	GAAATCGCCA	ACGAAGCCCA	TGGCTCCTCC
TTTCGGCTAG	CTGGCGCGGT	GTAGCAGTGT	CTTTAGCGGT	TGCTTCGGGT	ACCGAGGAGG
1870	1880	1890	1900	1910	1920
AAGGTTGCGG	CCATTGACCC	GGCCGCTGCC	AAACTCGCCG	ACTGCGGCGA	GTCTCGCAAG
TTCCAACGCC	GGTAACTGGG	CCGGCGACGG	TTTGAGCGGC	TGACGCCGCT	CAGAGCGTTC
1930	1940	1950	1960	1970	1980
CCGGGCGGGC	GCACCCGCGA	GGGCCGCGCA	CACTTTTCCC	AGACCTTTCA	TAGGCCGTCT
GGCCCGCCCG	CGTGGGCGCT	CCCGGCGCGT	GTGAAAAGGG	TCTGGAAAGT	ATCCGGCAGA
1990	2000	2010	2020	2030	2040
GCGACCGCTC	GCGGATCGAG	ACGGCGACGC	CGATTGGCGC	AAATGTCGTT	GCCCGAATTT
CGCTGGCGAG	CGCCTAGCTC	TGCCGCTGCG	GCTAACC GCG	TTTACAGCAA	CGGGCTTAAA
2050	2060	2070	2080	2090	2100
TCGGCGCCCT	CTATGAGGGG	CGTAGATAGA	GCTTCACGAT	GATGCAAGGA	TTCTCTCCAT
AGCCGCGGGA	GATACTCCCC	GCATCTATCT	CGAAGTGCTA	CTACGTTCTT	AAGGAGGGTA
2110	2120	2130	2140	2150	2160
GAGTGCCAGC	GGCCTGCCGT	TTGATGATTT	TCGCGAATTG	TTGCGCAACC	TGCCGGGGCC
CTCACGGTGC	CCGGACGGCA	AACTACTAAA	AGCGCTTAAC	AACGCGTTGG	ACGGCCCCGG
2170	2180	2190	2200	2210	2220
GGATGCGGCA	GCCCTCGTTG	CCGCGCGGGA	GCGGGACGCC	CAGCTGACGA	AGCCGCCGGG
CCTACGCCGT	CGGGAGCAAC	GGCGCGCCCT	CGCCCTGCGG	GTCGACTGCT	TCGGCGGGCC
2230	2240	2250	2260	2270	2280
CGCGCTCGGC	CGCCTCGAGG	AAATCGCCTT	CTGGCTCGCC	GCCTGGACGG	GCAAGGCGCC
GCGCGAGCCG	GCGGAGCTCC	TTTAGCGGAA	GACCGAGCGG	CGGACCTGCC	CGTTCCGCGG
2290	2300	2310	2320	2330	2340
GGTGGTCAAC	CGGCCGCTGG	TGGCGATCTT	TGCCGGCAAC	CACGGCGTCA	CCCGCCAGGG
CCACCAGTTG	GCCGGCGACC	ACC GCTAGAA	ACGGCCGTTG	GTGCCGCACT	GGGCGGTCCC
2350	2360	2370	2380	2390	2400
GGTGACCCCG	TTCCCGTCAT	CCGTCACCGC	ACAGATGGTC	GAGAATTTTG	CCGCCGGTGG
CCACTGGGGC	AAGGGCAGTA	GGCAGTGGCG	TGTCTACCAG	CTCTTAAAAC	GGCGGCCACC

FIG. 33D

2410	2420	2430	2440	2450	2460
CGCTGCGATC	AACCAGATCT	GCGTCAGCCA	CGACCTCGGG	CTGAAGGTCT	TCGACCTCGC
GCGACGCTAG	TTGGTCTAGA	CGCAGTCGGT	GCTGGAGCCC	GACTTCCAGA	AGCTGGAGCG
2470	2480	2490	2500	2510	2520
ACTCGAATAC	CCGACCGGTG	ATATCACCGA	GGAAGCCGCG	CTGTCCGAGC	GCGATTGCGC
TGAGCTTATG	GGCTGGCCAC	TATAGTGGCT	CCTTCGGCGC	GACAGGCTCG	CGCTAACGCG
2530	2540	2550	2560	2570	2580
CGCGACCATG	GCCTTTGGCA	TGGAGGCGAT	TGCCGGCGGC	ACGGATCTTC	TGTGCATCGG
GCGCTGGTAC	CGGAAACCGT	ACCTCCGCTA	ACGGCCGCCG	TGCCTAGAAG	ACACGTAGCC
2590	2600	2610	2620	2630	2640
CGAAATGGGC	ATCGGCAACA	CCACGATCGC	GGCCGCGATC	AATCTCGGCC	TTTATGGTGG
GCTTTACCCG	TAGCCGTTGT	GGTGCTAGCG	CCGGCGCTAG	TTAGAGCCGG	AAATACCACC
2650	2660	2670	2680	2690	2700
CACGGCCGAA	GAATGGGTCT	GTCCGGGTAC	CGGCTCCGAG	GGCGAGGTGC	TGAAGCGCAA
GTGCCGGCTT	CTTACCCAGC	CAGGCCCATG	GCCGAGGCTC	CCGCTCCACG	ACTTCGCGTT
2710	2720	2730	2740	2750	2760
GATCGCCGCG	GTCGAAAAGG	CCGTGGCGCT	GCATCGCGAT	CACCTGTCCG	ATCCGCTCGA
CTAGCGGCGC	CAGCTTTTCC	GGCACC CGCA	CGTAGCGCTA	GTGGACAGGC	TAGGCGAGCT
2770	2780	2790	2800	2810	2820
ACTGATGCGT	CGCCTCGGCG	GTCGTGAGAT	CGCGGCCATG	GCTGGCGCCA	TCCTGGCCGC
TGACTACGCA	GCGGAGCCGC	CAGCACTCTA	GCGCCGGTAC	CGACCGCGGT	AGGACCGGCG
2830	2840	2850	2860	2870	2880
CCGCGTCCAG	AAGGTACCTG	TCATCATCGA	CGGCTACGTG	GCGACCGCTG	CGGCTTCGAT
GGCGCAGGTC	TTCCATGGAC	AGTAGTAGCT	GCCGATGCAC	CGCTGGCGAC	GCCGAAGCTA
2890	2900	2910	2920	2930	2940
CCTGAAGGCG	GCCAACCCGT	CGGCCCTCGA	CCATTGCCTG	ATCGGCCATG	TTTCGGGCGA
GGACTTCCGC	CGGTTGGGCA	GCCGGGAGCT	GGTAACGGAC	TAGCCGGTAC	AAAGCCCGCT
2950	2960	2970	2980	2990	3000
ACCGGGGCAT	CTGCGCGCGA	TCGAGAAGCT	CGGCAAGACG	CCGCTGCTGG	CACTCGGCAT
TGCCCCCGTA	GACGCGCGCT	AGCTCTTCGA	GCCGTTCTGC	GGCGACGACC	GTGAGCCGTA

FIG. 33E

3010	3020	3030	3040	3050	3060
GCGGCTTGGC	GAAGGCACGG	GCGCGGCCCT	TGCCGCCGGT	ATCGTCAAGG	CGGCGGCCGC
CGCCGAACCG	CTTCCGTGCC	C GCGCCGGGA	ACGGCGGCCA	TAGCAGTTCC	GCCGCCGGCG
3070	3080	3090	3100	3110	3120
TTGCCACAGC	GGCATGGCGA	CCTTTGCCCA	GGCCGGCGTC	AGCAACAAGG	AATAGTGAAG
AACGGTGTGC	CCGTACCGCT	GGAAACGGGT	CCGGCCGCAG	TCGTTGTTCC	TTATCACTTC
3130	3140	3150	3160	3170	3180
TTCCGGCCGG	GCTTTGCAGG	AAGGCCGGCC	GGTTTCTGTC	CAAGGCCTGT	CACGGGCGCG
AAGGCCGGCC	CGAAACGTCC	TTCCGGCCGG	CCAAAGACAG	GTTCCGGACA	GTGCCCCGCG
3190	3200	3210	3220	3230	3240
AAGCTGTTCG	GTGCCGGGCC	TTGATGGATG	CGTCCTTCTC	GCCTATCCAA	AGCGCAAATG
TTCGACAGCG	CACGGCCCGG	AACTACCTAC	GCAGGAAGAG	CGGATAGGTT	TCGCGTTTAC
3250	3260	3270	3280	3290	3300
CGCGCCCTAG	CTATAGTCTT	GGGTGCCTGC	AACCGAGACC	GCCTTGCAAT	CGCCTCAATC
GCGCGGGATC	GATATCAGAA	CCCACGGACG	TTGGCTCTGG	CGGAACGTAA	GCGGAGTTAG
3310	3320	3330	3340	3350	3360
ACGATGTCGA	AGCAAGCACA	GTTTCAAGCC	CTGTCGAGAC	GAAATGGACG	CCAAGAACAC
TGCTACAGCT	TCGTTCGTGT	CAAAGTTCGG	GACAGCTCTG	CTTTACCTGC	GGTTCTTGTC
3370	3380	3390	3400	3410	3420
CACGCACCGC	ATTGGACAGA	CGGGTCCTGT	CGAGAAGCAG	ACCGGCATTC	GGCATCTCTT
GTGCGTGGCG	TAACCTGTCT	GCCCAGGACA	GCTCTTCGTC	TGGCCGTAAG	CCGTAGAGAA
3430	3440	3450	3460	3470	3480
TGCCGCTGCG	AGCTATTTCG	TCGGCGGGCG	CAAGCGGCTG	ATCGGCGAGG	CTGCCTTTTC
ACGGCGACGC	TCGATAAGCG	AGCCGCCGCG	GTTCGCCGAC	TAGCCGCTCC	GACGGAAAGC
3490	3500	3510	3520	3530	3540
CCACGAGCTG	ATCGCCTTTG	CCGCCGCGAT	GATCGCTTTC	ATCATCGTCG	GCGCAACCTT
GGTGCTCGAC	TAGCGGAAAC	GGCGGCGCTA	CTAGCGAAAG	TAGTAGCAGC	CGCGTTGGAA
3550	3560	3570	3580	3590	3600
CTTCCAATAT	GTGGCGATGG	CGATCCTGTT	CCTGCTGATG	ATGGCCTTCG	AGGCGATCAA
GAAGGTTATA	CACCGCTACC	GCTAGGACAA	GGACGACTAC	TACCGGAAGC	TCCGCTAGTT

FIG. 33F

3610	3620	3630	3640	3650	3660
CACGGCAATC	GAGGAAATTG	TCGATCGCGT	TTCTCCCGAA	ATCTCGGAAA	TGGGTAAGAA
GTGCCGTTAG	CTCCTTTAAC	AGCTAGCGCA	AAGAGGGCTT	TAGAGCCTTT	ACCCATTCTT
3670	3680	3690	3700	3710	3720
CGCCAAGGAT	CTCGGCTCCT	TCGCCTGCCT	CTGCCTGATT	GTCGCCAACG	GTGTCTATGC
GCGGTTCCCTA	GAGCCGAGGA	AGCGGACGGA	GACGGACTAA	CAGCGGTTGC	CACAGATACG
3730	3740	3750	3760	3770	3780
CGCCTATGTC	GTGATCTTCG	ACGGCTTCAT	GAAGTACCGG	GCTAGCGGGC	CGGCGCCTTC
GCGGATACAG	CACTAGAAGC	TGCCGAAGTA	CTTGACTGGC	CGATCGCCCG	GCCGCGGAAG
3790	3800	3810	3820	3830	3840
ACCCGATAAA	GCACATGCGG	ACGCAGCGGG	TTGCCCCCGG	GTACCGTGAC	GTCGTCGAAA
TGGGCTATTT	CGTGTACGCC	TGCGTCGCCC	AACGGGGGCC	CATGGCACTG	CAGCAGCTTT
3850	3860	3870	3880	3890	3900
TCATCAGCCG	GATCC				
AGTAGTCGGC	CTAGG				

FIG. 33G

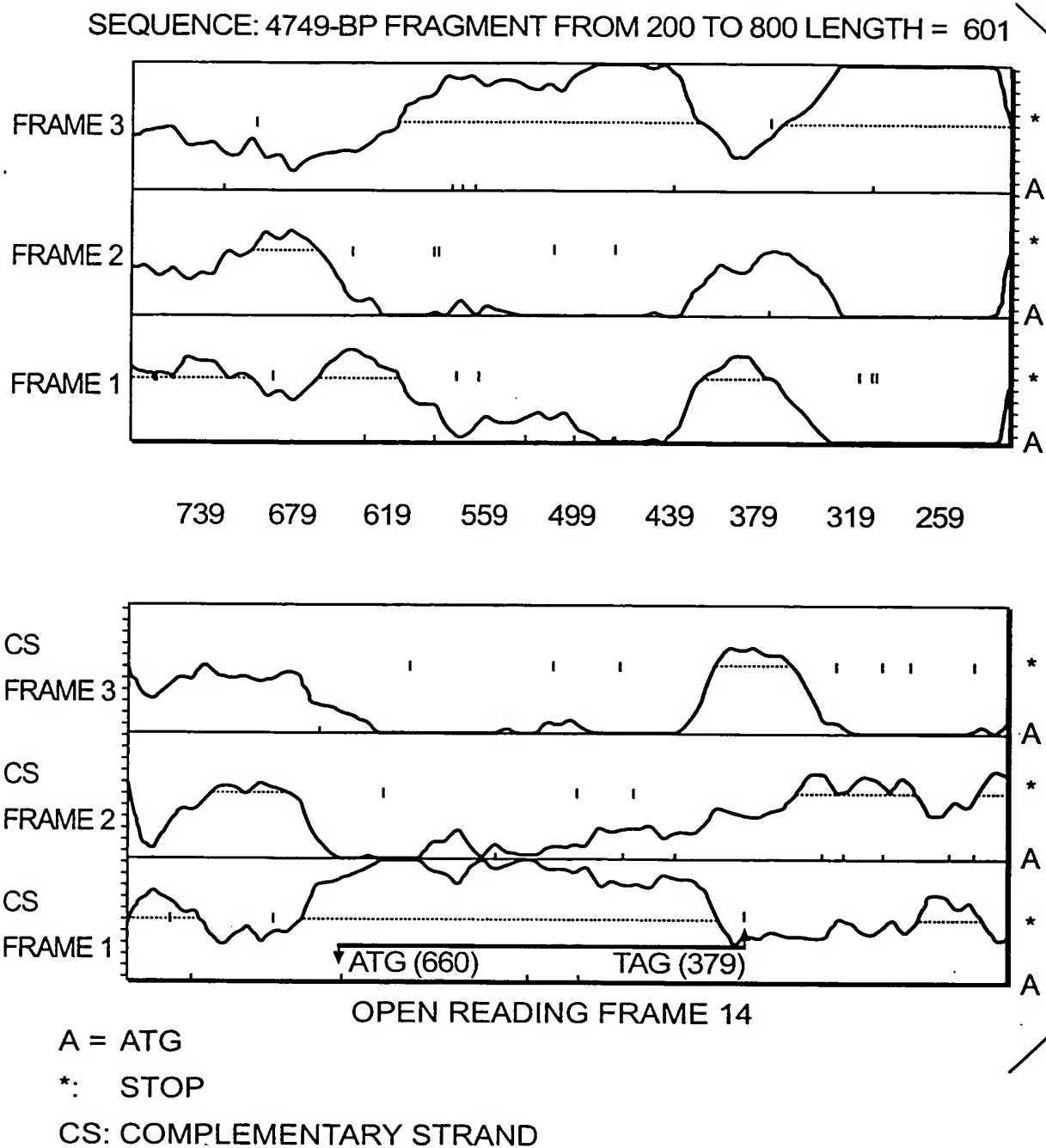


FIG. 34A

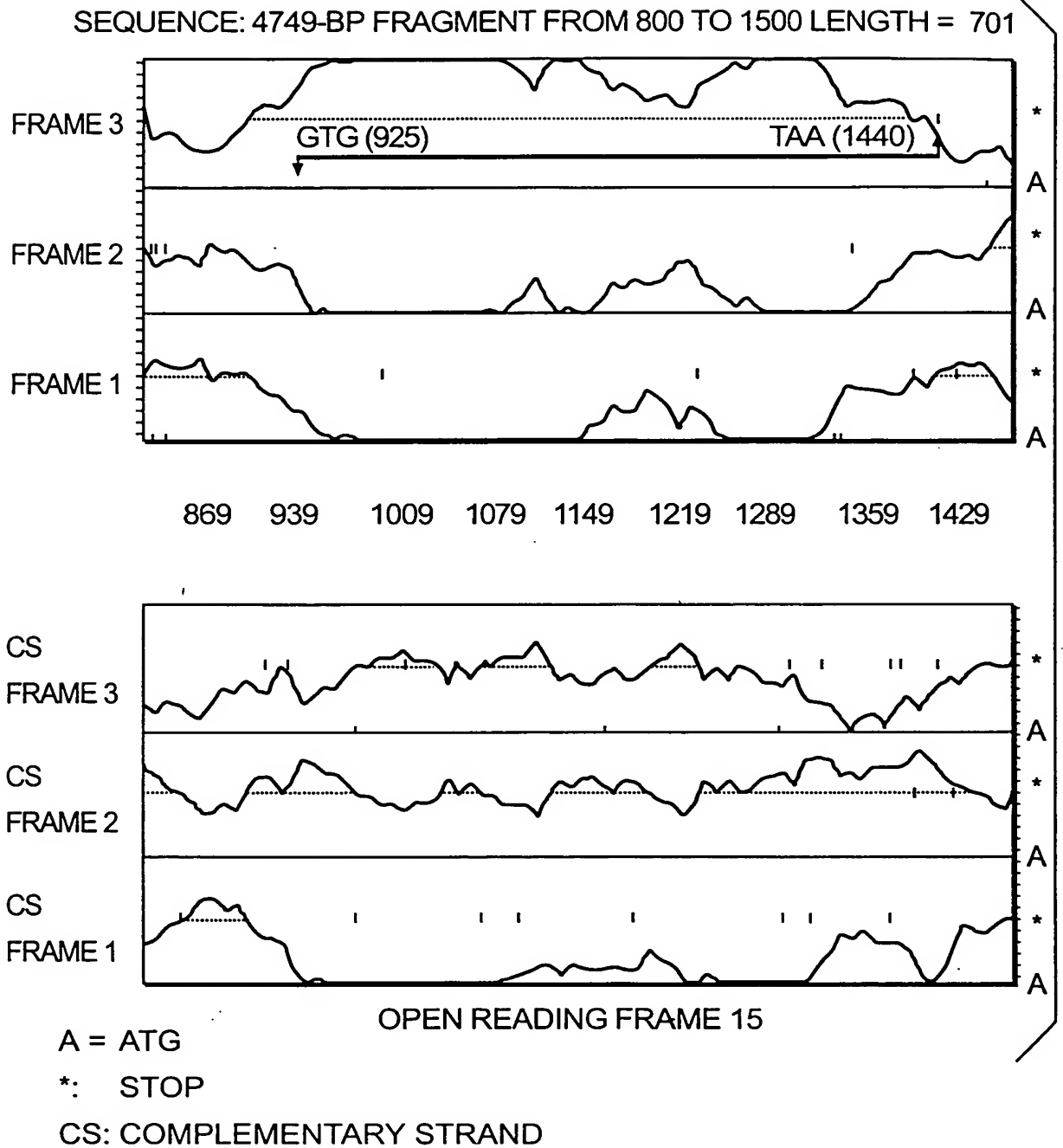
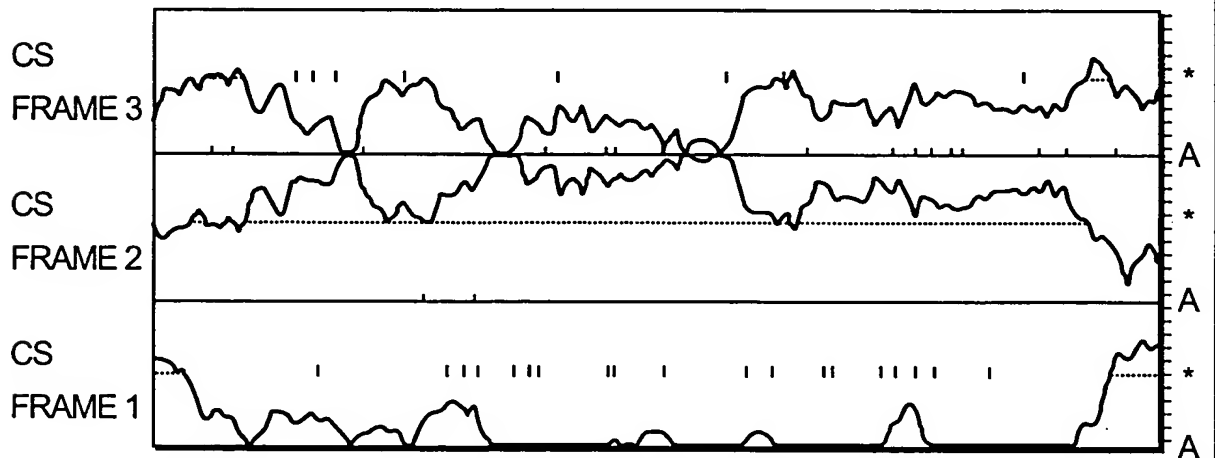
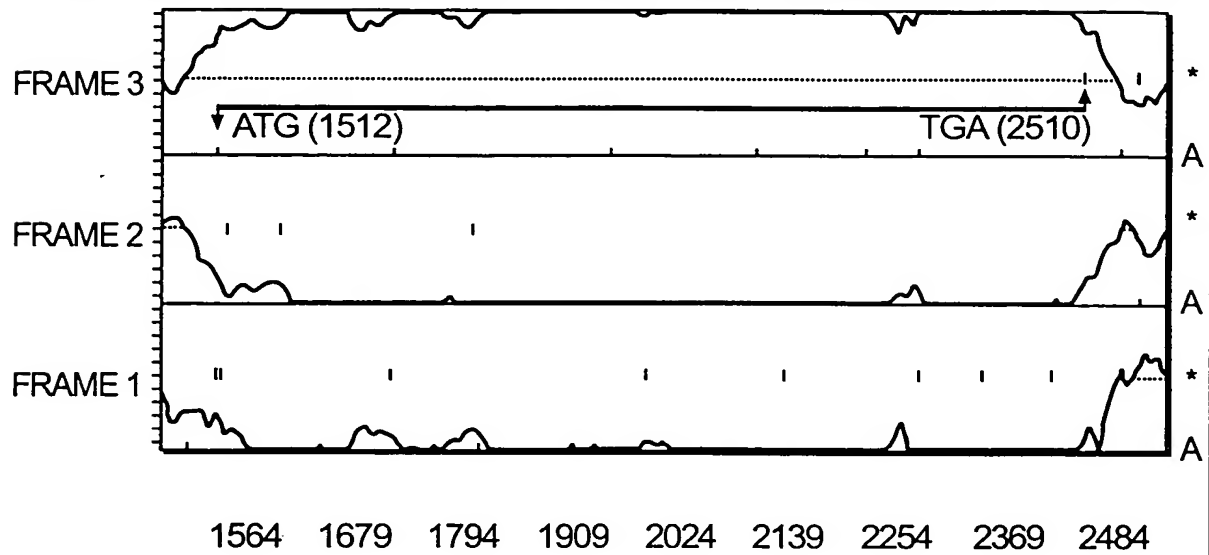


FIG. 34B

SEQUENCE: 4749-BP FRAGMENT FROM 1450 TO 2600 LENGTH = 1151 BP



A = ATG

*: STOP

CS: COMPLEMENTARY STRAND

OPEN READING FRAME 16

FIG. 34C

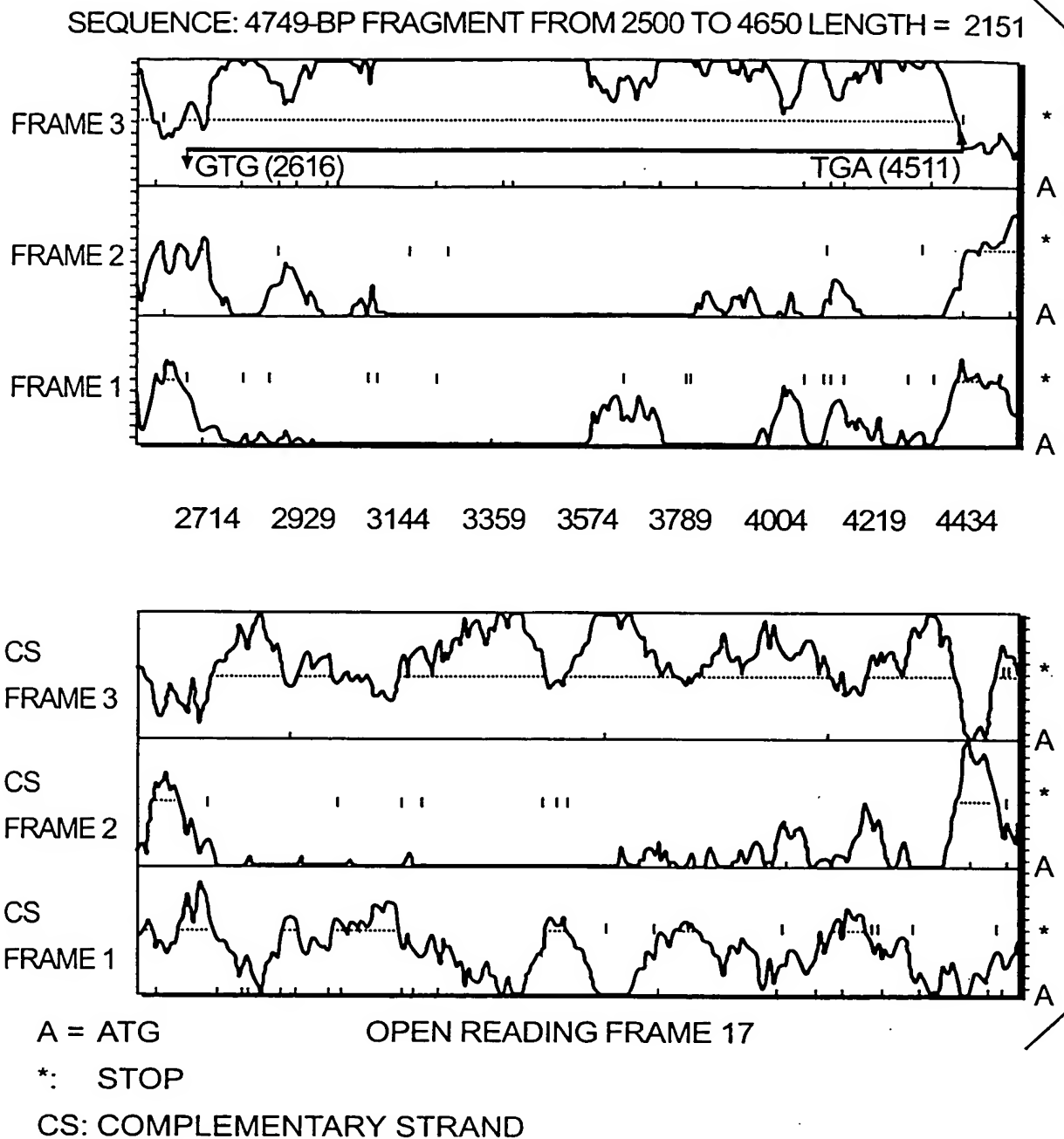


FIG. 34D

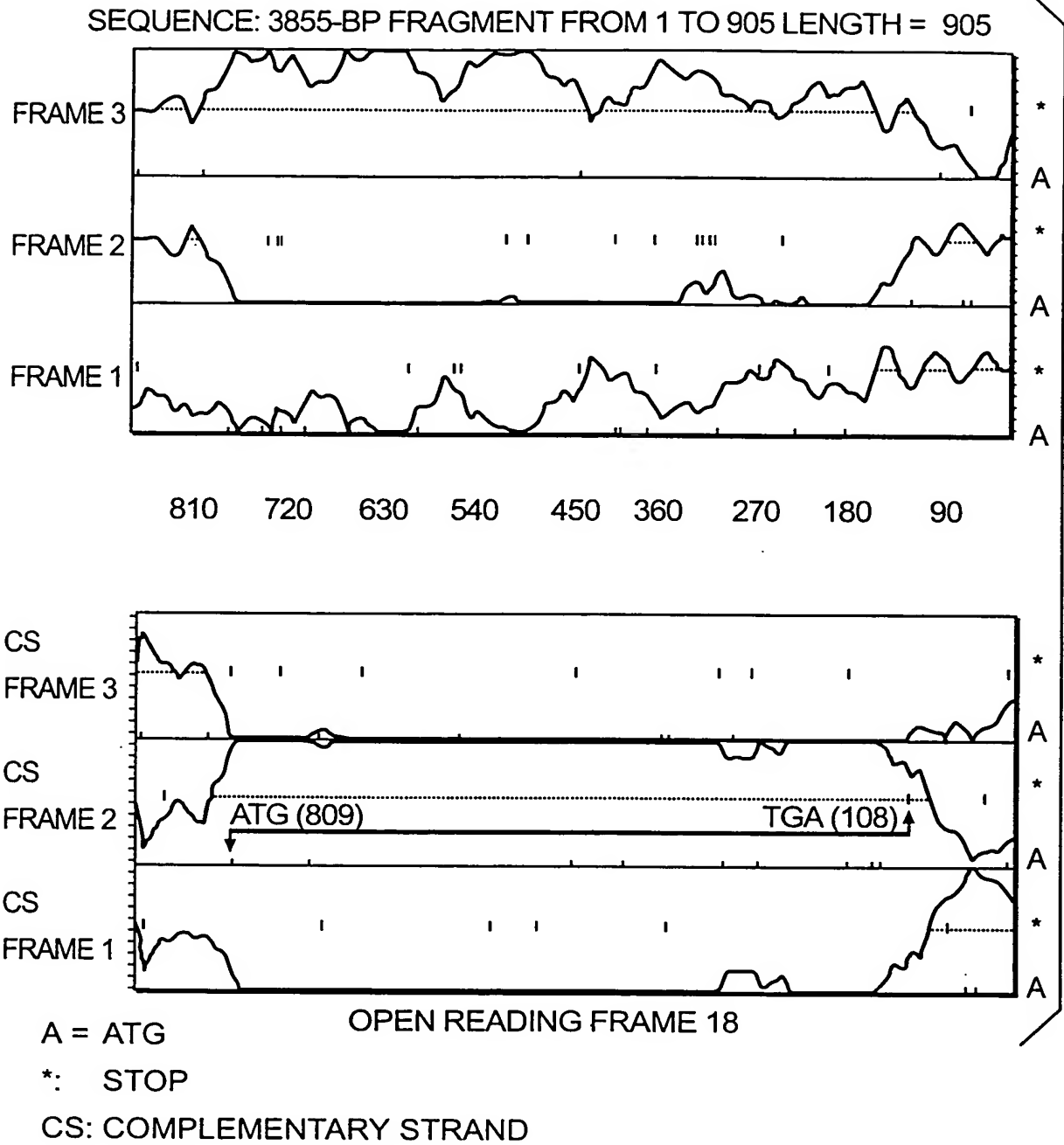
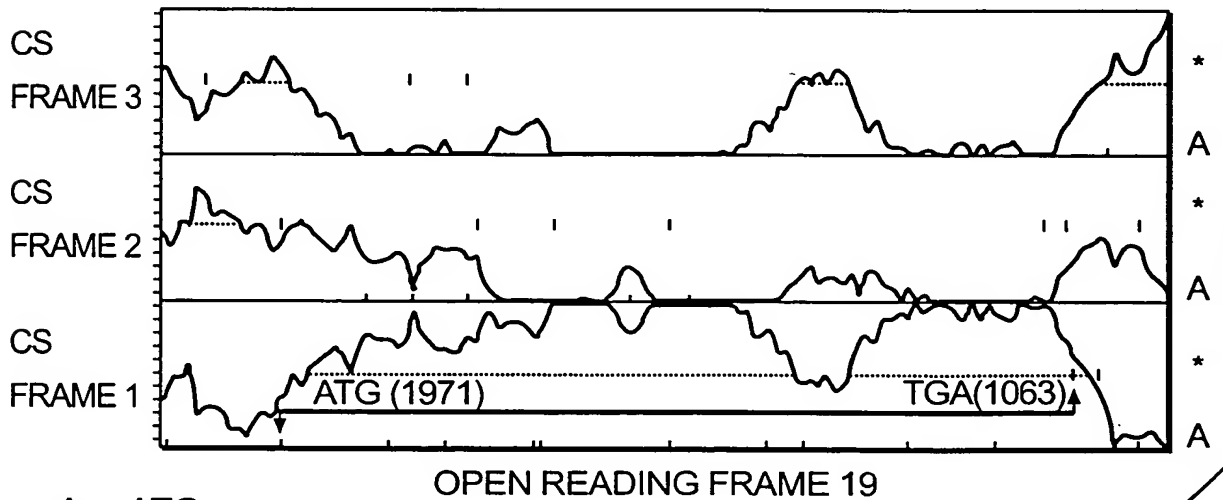
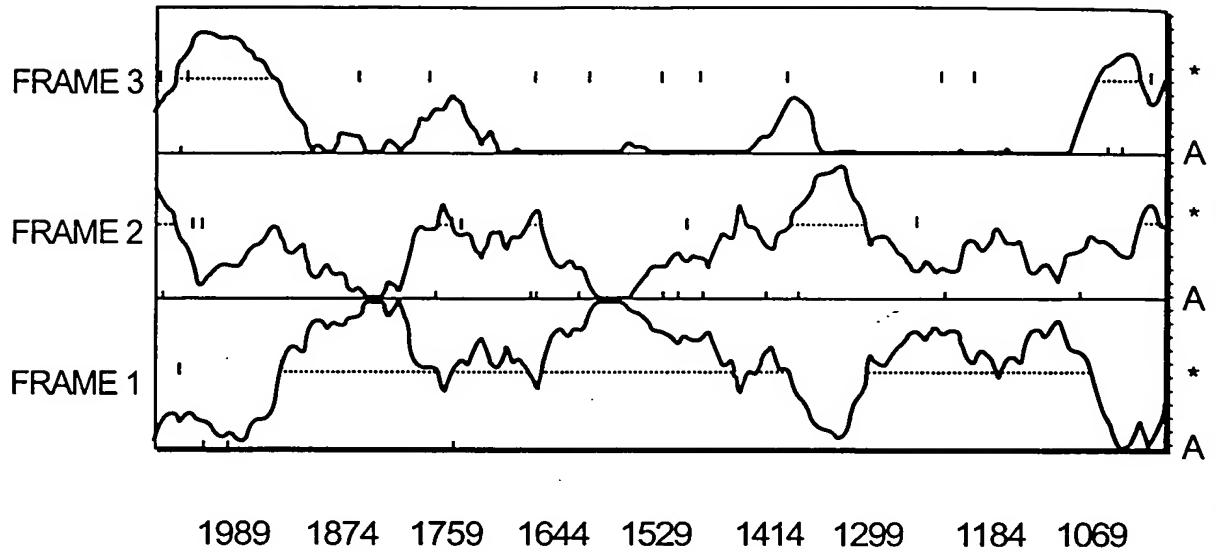


FIG. 35A

SEQUENCE: 3855-BP FRAGMENT FROM 955 TO 2105 LENGTH = 1151



A = ATG

*: STOP

CS: COMPLEMENTARY STRAND

FIG. 35B

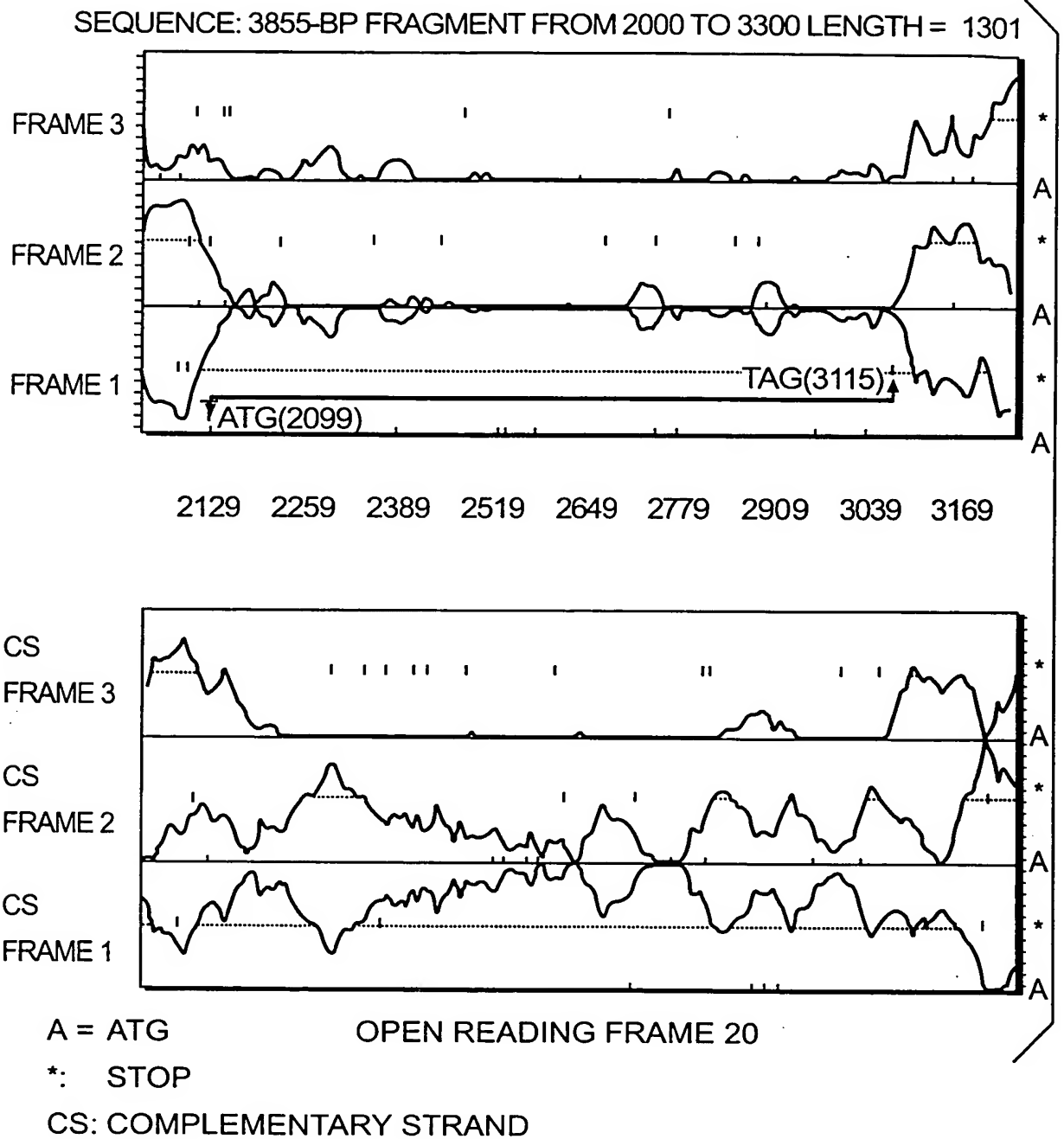
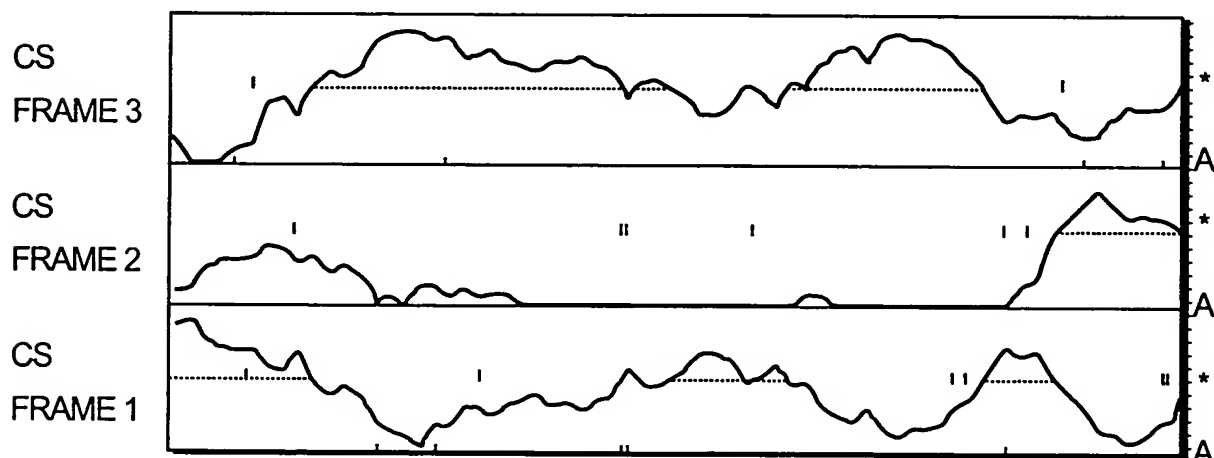
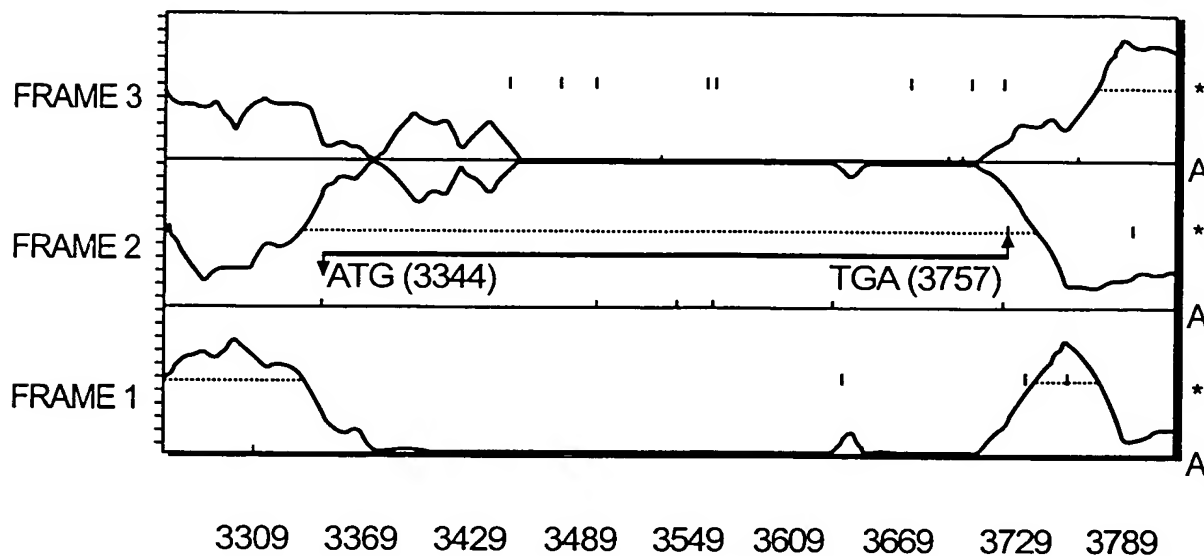


FIG. 35C

SEQUENCE: 3855-BP FRAGMENT FROM 3250 TO 3855 LENGTH = 606



A = ATG

OPEN READING FRAME 21

*: STOP

CS: COMPLEMENTARY STRAND

FIG. 35D



FIG. 36

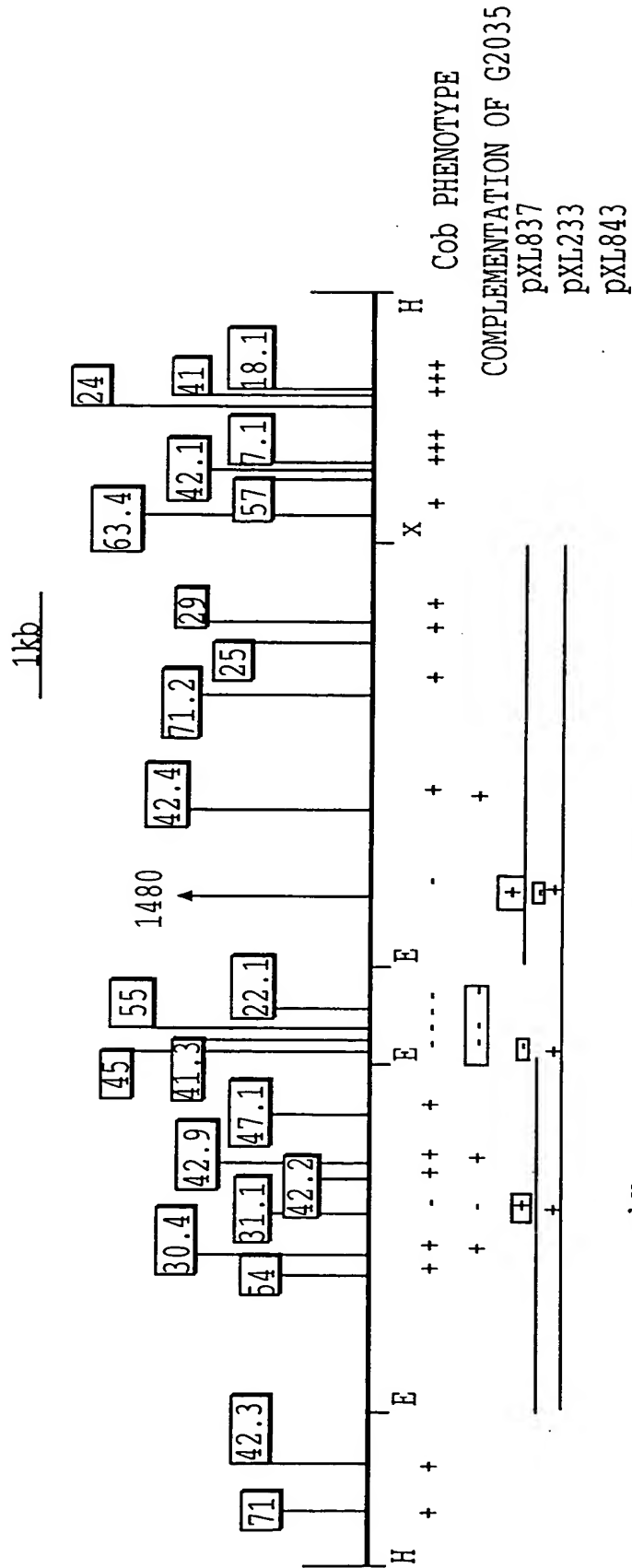


FIG. 37

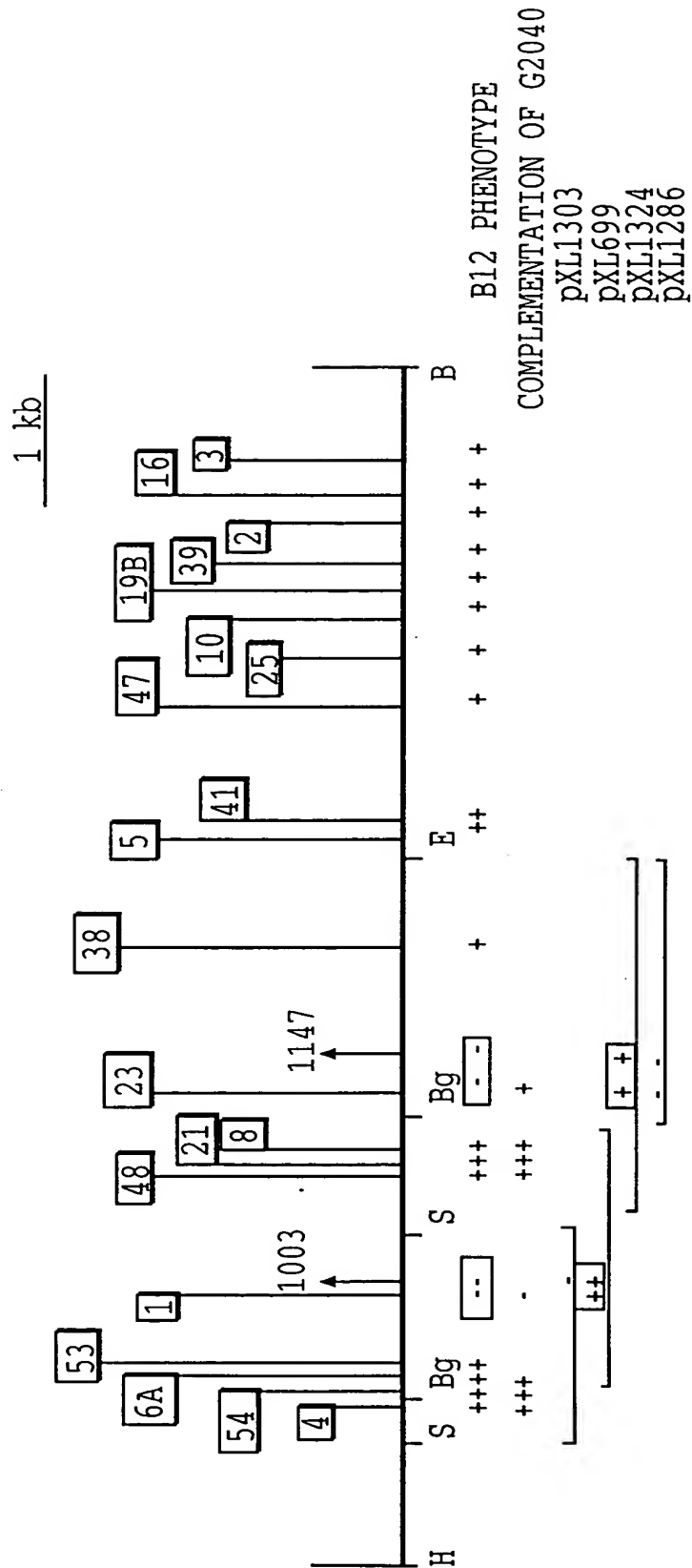


FIG. 39

ORF 18 ORF 19 ORF 20 ORF 21

NAME = COBS			FIRST RESIDUE = 1		LAST RESIDUE = 332	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	15	4.52	2206.03	5.97
2	LEU	L	29	8.73	3279.44	8.87
3	ILE	I	20	6.02	2261.68	6.12
4	MET	M	7	2.11	917.28	2.48
5	VAL	V	33	9.94	3269.26	8.85
6	SER	S	15	4.52	1305.48	3.53
7	PRO	P	11	3.31	1067.58	2.89
8	THR	T	24	7.23	2425.14	6.56
9	ALA	A	32	9.64	2273.19	6.15
10	TYR	Y	9	2.71	1467.57	3.97
11	*	*	0	0.00	0.00	0.00
12	HIS	H	10	3.01	1370.59	3.71
13	GLN	Q	10	3.01	1280.59	3.46
14	ASN	N	12	3.61	1368.52	3.70
15	LYS	K	13	3.92	1665.23	4.51
16	ASP	D	28	8.43	3220.75	8.71
17	GLU	E	15	4.52	1935.64	5.24
18	CYS	C	3	0.90	309.03	0.84
19	TRP	W	4	1.20	744.32	2.01
20	ARG	R	22	6.63	3434.22	9.29
21	GLY	G	20	6.02	1140.43	3.09
22	-	-	0	0.00	0.00	0.00

RESIDUES = 332
 MOLECULAR WEIGHT (MONOISOTOPIC) = 36960.0000
 MOLECULAR WEIGHT (AVERAGE) = 36983.1797
 INDEX OF POLARITY (%) = 44.88
 ISOELECTRIC POINT = 6.34
 OD 260 (1 mg/ml) = 0.611 OD 280 (1 mg/ml) = 0.891

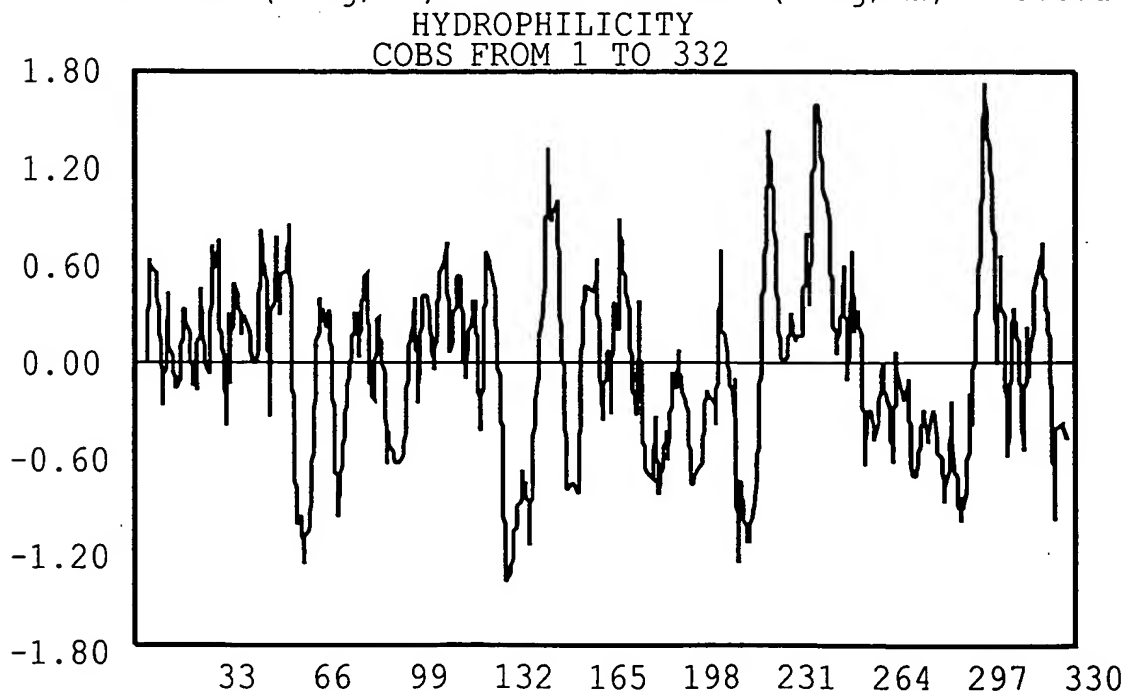


FIG. 40A

cobS GENE (SEQ ID NO: 31) AND COBS PROTEIN (SEQ ID NO: 32)
 SEQUENCE OF THE 4749-BP SalI-SalI-SalI-SalI-SalI-BglI FRAGMENT
 FROM 1512 TO 2510

```

MetMetSerLysIleAspLeuAspIleSerAsnLeuProAspThrThrIleSerValArgGluValPheGlyIle
ATGATGAGCAAGATTGACCTCGACATTTCCAACCTCCCCGACACCACGATTTCCGTCCGGGAGGTTTTTCGGTATT
1521      1531      1541      1551      1561      1571      1581
AspThrAspLeuArgValProAlaTyrSerLysGlyAspAlaTyrValProAspLeuAspProAspTyrLeuPhe
GATACGGATTTGCGCGTTTCCTGCCTATTTCGAAGGGCGACGCTATGTCCCGGATCTGGATCCGGACTACCTCTTC
1596      1606      1616      1626      1636      1646      1656
AspArgGluThrThrLeuAlaIleLeuAlaGlyPheAlaHisAsnArgArgValMetValSerGlyTyrHisGly
GACCGCGAAACGACGCTCGCCATTCTCGCAGGCTTCGCCCACAACCGACGCGTGATGGTGTGGGGCTATCACGGC
1671      1681      1691      1701      1711      1721      1731
ThrGlyLysSerThrHisIleGluGlnValAlaAlaArgLeuAsnTrpProCysValArgValAsnLeuAspSer
ACCGGCAAGTCCACCCATATCGAGCAGGTGCGCGCGCCTCAACTGGCCGTGCGTGCAGTCAACCTCGATAGC
1746      1756      1766      1776      1786      1796      1806
HisValSerArgIleAspLeuValGlyLysAspAlaIleValValLysAspGlyLeuGlnValThrGluPheLys
CATGTCAGCCGTATCGACCTCGTCGGCAAGGACGCGATCGTCGTCAGGACGGCTGCAGGTCACCGAATCAAG
1821      1831      1841      1851      1861      1871      1881
AspGlyIleLeuProTrpAlaTyrGlnHisAsnValAlaLeuValPheAspGluTyrAspAlaGlyArgProAsp
GACGGCATCCTGCCCTGGGCTACCAGCACAATGTGCGCGTCTGTCGACGAATACGATGCCGGCCGCCGGAC
1896      1906      1916      1926      1936      1946      1956
ValMetPheValIleGlnArgValLeuGluSerSerGlyArgLeuThrLeuLeuAspGlnSerArgValIleArg
GTCATGTTTCGTCATCCAGCGCGTGTGAATCCTCCGGCCGCTGACGCTGCTCGACCAGAGCCGTGTCATCCGT
1971      1981      1991      2001      2011      2021      2031
ProHisProAlaPheArgLeuPheAlaThrAlaAsnThrValGlyLeuGlyAspThrThrGlyLeuTyrHisGly
CCGCACCCGGCCTTCCGCCTGTTTGCACCGCCAACACCGTCGGCCTCGGCGACACGACCGGCCTCTATCACGGC
2046      2056      2066      2076      2086      2096      2106
ThrGlnGlnIleAsnGlnAlaGlnMetAspArgTrpSerIleValThrThrLeuAsnTyrLeuProHisAspLys
ACGCAGCAGATCAACCAGGCGCAGATGGACCGCTGGTCGATCGTCACCACGCTGAACCTACCTGCCGCACGACAAG
2121      2131      2141      2151      2161      2171      2181
GluValAspIleValAlaAlaLysValLysGlyPheThrAlaAspLysGlyArgGluThrValSerLysMetVal
GAAGTCGACATCGTCGCCGCAAGGTCAAGGGCTTACCGCCGACAAGGGCCGAGACCGTCTCCAAGATGGTA
2196      2206      2216      2226      2236      2246      2256
ArgValAlaAspLeuThrArgAlaAlaPheIleAsnGlyAspLeuSerThrValMetSerProArgThrValIle
CGTGTGCGCCGACCTACGCGCGCAGCCTTCATCAATGGCGATCTCTCGACTGTATGAGCCCGGTACGGTCATC
2271      2281      2291      2301      2311      2321      2331
ThrTrpAlaGluAsnAlaHisIlePheGlyAspIleAlaPheAlaPheArgValThrPheLeuAsnLysCysAsp
ACCTGGGCGCGAGAACGCCACATCTTCGGCGACATCGCTTTTCGCCTCCGCGTGACCTTCTCAACAAGTGCGAC
2346      2356      2366      2376      2386      2396      2406
GluLeuGluArgAlaLeuValAlaGluHisTyrGlnArgAlaPheGlyIleGluLeuLysGluCysAlaAlaAsn
GAGCTGGAGCGGGCGTGGTCGCCGAGCACTACCAGCGCGCTTCGGCATCGAGTGAAGGAATGCGCTGCCAAG
2421      2431      2441      2451      2461      2471      2481
IleValLeuGluAlaThrAla***
ATCGTGCTCGAAGCCACCGCCTGA
2496      2506

```

FIG. 40B

NAME = COBT			FIRST RESIDUE = 1		LAST RESIDUE = 631	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	16	2.54	2353.09	3.35
2	LEU	L	56	8.87	6332.71	9.01
3	ILE	I	29	4.60	3279.44	4.67
4	MET	M	18	2.85	2358.73	3.36
5	VAL	V	31	4.91	3071.12	4.37
6	SER	S	33	5.23	2872.06	4.09
7	PRO	P	24	3.80	2329.27	3.31
8	THR	T	28	4.44	2829.34	4.03
9	ALA	A	75	11.89	5327.78	7.58
10	TYR	Y	8	1.27	1304.51	1.86
11	*	*	0	0.00	0.00	0.00
12	HIS	H	7	1.11	959.41	1.36
13	GLN	Q	29	4.60	3713.70	5.28
14	ASN	N	22	3.49	2508.94	3.57
15	LYS	K	25	3.96	3202.37	4.56
16	ASP	D	49	7.77	5636.32	8.02
17	GLU	E	67	10.62	8645.85	12.30
18	CYS	C	5	0.79	515.05	0.73
19	TRP	W	6	0.95	1116.48	1.59
20	ARG	R	61	9.67	9522.17	13.55
21	GLY	G	42	6.66	2394.90	3.41
22	-	-	0	0.00	0.00	0.00

RESIDUES = 631
 MOLECULAR WEIGHT (MONOISOTOPIC) = 70291.3984
 MOLECULAR WEIGHT (AVERAGE) = 70334.7656
 INDEX OF POLARITY (%) = 50.87
 ISOELECTRIC POINT = 5.10
 OD 260 (1mg/ml) = 0.423 OD 280 (1mg/ml) = 0.610

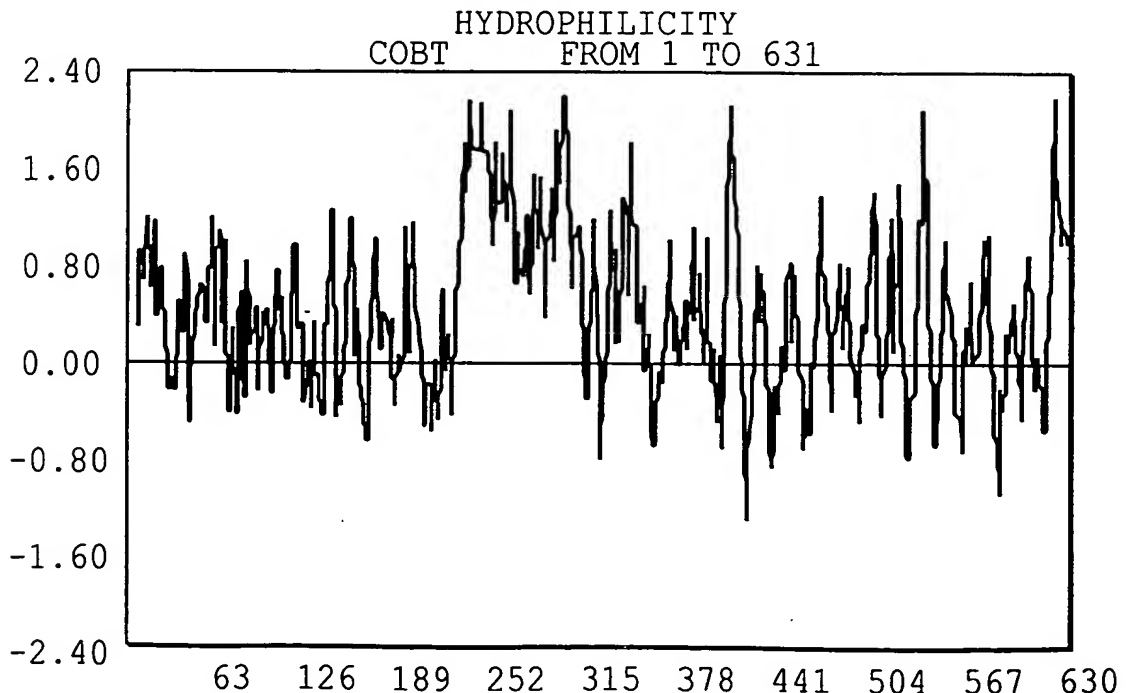


FIG. 40C

cobT GENE (SEQ ID NO: 33) AND COBT PROTEIN (SEQ ID NO: 44)
SEQUENCE OF THE 4749-BP SalI-SalI-SalI-SalI-SalI-BglI FRAGMENT
FROM 2616 TO 4511

ValSerSerAsnSerLysAlaLysProThrThrArgGluAsnAlaAlaGluProPheLysArgAlaLeuSerGly
GTGAGCTCGAATTCGAAGGCAAAGCCAACCACGCGCAGAATGCTGCGGAACCGTTCAAGCGGGCGCTTTCCGGC
2625 2635 2645 2655 2665 2675 2685
CysIleArgSerIleAlaGlyAspAlaGluValGluValAlaPheAlaAsnGluArgProGlyMetThrGlyGlu
TGCATCCGATCGATCGCGGGCGATGCCGAGGTGGAAGTCGCCTTCGCCAACGAGCGCGCGGCATGACCGGCGAA
2700 2710 2720 2730 2740 2750 2760
ArgIleArgLeuProGluLeuSerLysArgProThrLeuGlnGluLeuAlaValThrArgGlyLeuGlyAspSer
CGCATCCGTCTGCCGGAACCTTCCAAGCGCCCGACCCTGCAGGAACCTGCCGTGACCCGCGGGCTCGGTGACAGC
2775 2785 2795 2805 2815 2825 2835
MetAlaLeuArgLysAlaCysThrHisAlaArgIleGlnArgThrMetSerProGlnGlyAlaAspAlaArgAla
ATGGCGCTGCGCAAGGCCTGTACGCATGCGCGGATCCAGCGCACCATGTGCGCCGAAGGGGCGGACGCGCGCGC
2850 2860 2870 2880 2890 2900 2910
IlePheAspAlaValGluGlnAlaArgValGluAlaIleGlySerLeuArgMetAlaGlyValAlaLysAsnLeu
ATCTTCGATGCGGTGGAGCAGGCTCGTGTGAGGCGATCGGGTCGTTGCGCATGGCGGGTGTGCCAAGAACCTC
2925 2935 2945 2955 2965 2975 2985
AsnValMetLeuGluGluLysTyrAlaLysAlaAsnPheAlaThrIleGluArgGlnAlaAspAlaProLeuGly
AACGTCATGCTCGAAGAGAAATACGCCAAGGCGAATTCGCAACGATCGAGCGCCAGGCGGACGCGCCGCTCGGC
3000 3010 3020 3030 3040 3050 3060
GluAlaValAlaLeuLeuValArgGluLysLeuThrGlyGlnLysProProAlaSerAlaGlyLysValLeuAsp
GAGGCCGTAGCGCTGCTGGTGC GCGAGAAGCTGACGGGCCAGAAGCCGCCGGCGTCTGCCGGAAGGTGCTCGAC
3075 3085 3095 3105 3115 3125 3135
LeuTrpArgGluPheIleGluGlyLysAlaAlaGlyAspIleGluHisLeuSerSerThrIleAsnAsnGlnGln
CTCTGGCGCGAGTTTCATCGAGGGCAAGGCTGCCGGCGACATTGAGCACCTGTCGTCGACGATCAACAACCAGCAG
3150 3160 3170 3180 3190 3200 3210
AlaPheAlaArgValValArgAspMetLeuThrSerMetGluValAlaGluLysTyrGlyAspAspAspAsnGlu
GCCTTTGCCCCGGGTCGTTGCGGACATGCTGACCTCGATGGAAGTCGCCGAGAAATACGGTGACGACGACAACGAG
3225 3235 3245 3255 3265 3275 3285
ProAspGluGlnGluSerGluThrAspGluAspGlnProArgSerGlnGluGlnAspGluAsnAlaSerAspGlu
CCGGACGAGCAGGAAAGCGAGACCGACGAAGACCAGCCGCGCAGCCAGGAGCAGGACGAGAACGCCAGCGACGAG
3300 3310 3320 3330 3340 3350 3360
GluAlaGlyAspAspAlaAlaProAlaAspGluAsnGlnAlaAlaGluGluGlnMetGluGluGlyGluMetAsp
GAAGCCGGCGACGATGCCGACCCGCCGACGAGAACCAGGCTGCCGAAGAGCAGATGGAAGAAGGCGAGATGGAC
3375 3385 3395 3405 3415 3425 3435
GlyAlaGluIleSerAspAspAspLeuGlnAspGluGlyAspGluAspSerGluThrProGlyGluValLysArg
GGCGCGGAGATCTCCGACGACGATCTCCAGGACGAAGGCGACGAGACGCGAAACGCCCGCGGAGGTCAAGCGT
3450 3460 3470 3480 3490 3500 3510

FIG. 40D

ProAsnGlnProPheAlaAspPheAsnGluLysValAspTyrAlaValPheThrArgGluPheAspGluThrIle
 CCGAACCAGCCCTTCGCCGACTTCAACGAGAAGGTCGACTACGCCGTCTTACCCGCGAGTTCGACGAGACGATT
 3525 3535 3545 3555 3565 3575 3585
 AlaSerGluGluLeuCysAspGluAlaGluLeuAspArgLeuArgAlaPheLeuAspLysGlnLeuAlaHisLeu
 GCCTCGAAGAGCTTTGCGACGAGGCCGAGCTCGACCGGCTGCGCGCCTTCCTCGACAAGCAGCTTGCCCATCTT
 3600 3610 3620 3630 3640 3650 3660
 GlnGlyAlaValGlyArgLeuAlaAsnArgLeuGlnArgArgLeuMetAlaGlnGlnAsnArgSerTrpGluPhe
 CAAGGCGCGGTCGGCCGCTTGCCAACCGGCTGCAGCGCCGCTGATGGCGCAGCAGAACCGCTCCTGGGAGTTC
 3675 3685 3695 3705 3715 3725 3735
 AspLeuGluGluGlyTyrLeuAspSerAlaArgLeuGlnArgIleIleIleAspProMetGlnProLeuSerPhe
 GATCTCGAAGAGGGGTATCTCGATTGCGCGCGGCTTCAGCGCATCATCATCGATCCGATGCAGCCGCTTTCCTTC
 3750 3760 3770 3780 3790 3800 3810
 LysArgGluLysAspThrAsnPheArgAspThrValValThrLeuLeuIleAspAsnSerGlySerMetArgGly
 AAGCGCGAAAAGGACACCAACTTCCGCGATACCGTCGTGACGCTGCTGATCGACAATTCCGGCTCGATGCGCGGC
 3825 3835 3845 3855 3865 3875 3885
 ArgProIleThrValAlaAlaThrCysAlaAspIleLeuAlaArgThrLeuGluArgCysGlyValLysValGlu
 CGTCCGATCACGGTTGCCGCCACCTGCGCCGATATCCTGGCGCGCACGCTCGAGCGCTGCGGCGTCAAGGTCGAG
 3900 3910 3920 3930 3940 3950 3960
 IleLeuGlyPheThrThrLysAlaTrpLysGlyGlyGlnSerArgGluLysTrpLeuAlaGlyGlyLysProGln
 ATCCTCGGTTTTACCACCAAGGCGTGGAAGGGTGGGCAGTCACGCGAGAAGTGGCTGGCCGGCGGCAAGCCACAG
 3975 3985 3995 4005 4015 4025 4035
 AlaProGlyArgLeuAsnAspLeuArgHisIleValTyrLysSerAlaAspAlaProTrpArgArgAlaArgArg
 GCCCCGGGTCGCCTCAACGACCTGCGACACATCGTCTACAAGTCTGCCGACGCTCCGTGGCGCCGGGCACGACGC
 4050 4060 4070 4080 4090 4100 4110
 AsnLeuGlyLeuMetMetArgGluGlyLeuLeuLysGluAsnIleAspGlyGluAlaLeuIleTrpAlaHisGlu
 AATCTCGGCCTGATGATGCGGGAAGGCCTGCTCAAGGAAAACATCGACGGCGAGGCGTTGATTTGGGCGCATGAG
 4125 4135 4145 4155 4165 4175 4185
 ArgLeuMetAlaArgArgGluGlnArgArgIleLeuMetMetIleSerAspGlyAlaProValAspAspSerThr
 CGGCTGATGGCGCGGCGCGAACAGCGGCGCATCTGATGATGATTTCCGACGGCGCGCCGTCGACGACTCGACG
 4200 4210 4220 4230 4240 4250 4260
 LeuSerValAsnProGlyAsnTyrLeuGluArgHisLeuArgAlaValIleGluGlnIleGluThrArgSerPro
 CTGTCCGTCAATCCAGGAACTATCTGGAGCGTCACCTGCGCGCGGTCATCGAGCAGATCGAAACGCGCTCGCCG
 4275 4285 4295 4305 4315 4325 4335
 ValGluLeuLeuAlaIleGlyIleGlyHisAspValThrArgTyrTyrArgArgAlaValThrIleValAspAla
 GTGGAAGTCTGGCGATCGGTATCGGCCACGACGTGACGCGCTACTATCGCCGTGCCGTACCATCGTCGATGCC
 4350 4360 4370 4380 4390 4400 4410
 AspGluLeuAlaGlyAlaMetThrGluGlnLeuAlaAlaLeuPheGluAspGluSerGlnArgArgGlySerSer
 GATGAGCTTGCCGGCGCGATGACCGAACAGCTGGCCGCACTCTTCGAGGACGAAAGCCAGCGCCGCGGTTCTTCG
 4425 4435 4445 4455 4465 4475 4485
 ArgLeuArgArgAlaGly***
 CGTCTTCGCCGCGCCGGGTGA
 4500 4510

FIG. 40E

NAME = COBX			FIRST RESIDUE = 1		LAST RESIDUE = 93	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	3	3.23	441.21	4.29
2	LEU	L	8	8.60	904.67	8.80
3	ILE	I	5	5.38	565.42	5.50
4	MET	M	3	3.23	393.12	3.82
5	VAL	V	5	5.38	495.34	4.82
6	SER	S	6	6.45	522.19	5.08
7	PRO	P	6	6.45	582.32	5.66
8	THR	T	5	5.38	505.24	4.92
9	ALA	A	10	10.75	710.37	6.91
10	TYR	Y	0	0.00	0.00	0.00
11	*	*	0	0.00	0.00	0.00
12	HIS	H	7	7.53	959.41	9.33
13	GLN	Q	2	2.15	256.12	2.49
14	ASN	N	2	2.15	228.09	2.22
15	LYS	K	3	3.23	384.28	3.74
16	ASP	D	3	3.23	345.08	3.36
17	GLU	E	10	10.75	1290.43	12.55
18	CYS	C	0	0.00	0.00	0.00
19	TRP	W	1	1.08	186.08	1.81
20	ARG	R	7	7.53	1092.71	10.63
21	GLY	G	7	7.53	399.15	3.88
22	-	-	0	0.00	0.00	0.00

RESIDUES = 93
 MOLECULAR WEIGHT (MONOISOTOPIC) = 10279.2354
 MOLECULAR WEIGHT (AVERAGE) = 10285.6309
 INDEX OF POLARITY (%) = 48.39
 ISOELECTRIC POINT = 6.94
 OD 260 (1mg/ml) = 0.411 OD 280 (1mg/ml) = 0.541

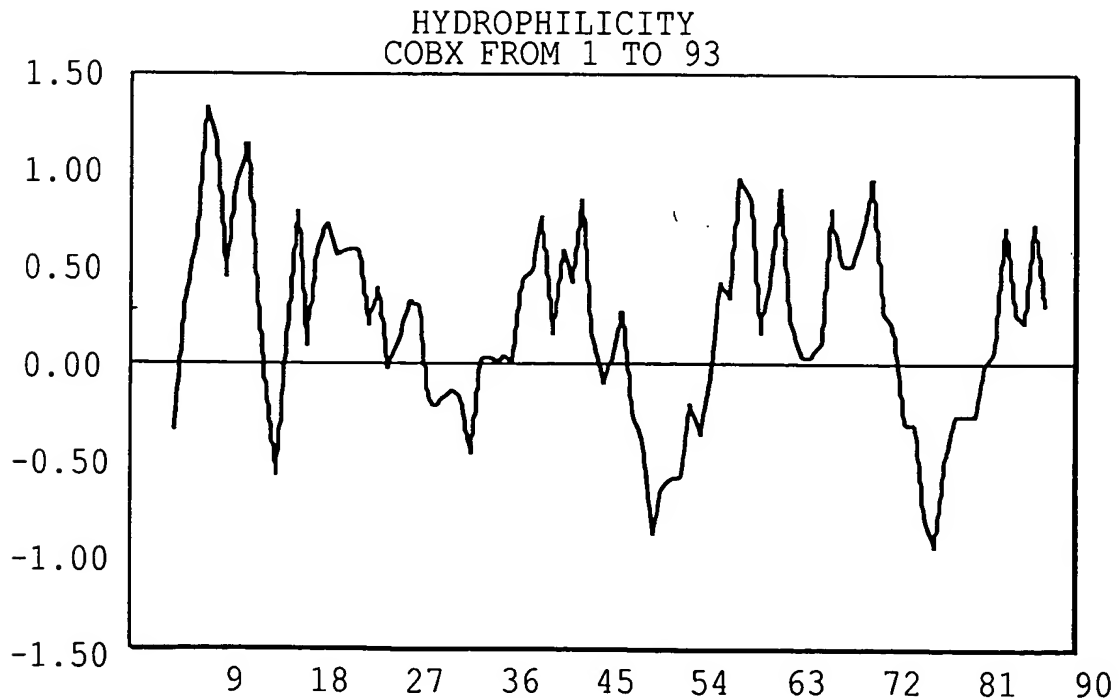


FIG. 40F

cobX GENE (SEQ ID NO: 35) AND COBX PROTEIN (SEQ ID NO: 36)
SEQUENCE OF THE 4749-BP BglI-SalI-SalI-SalI-SalI-SalI FRAGMENT
FROM 4089 TO 4370

```
MetSerLeuThrGluThrIleGluLysLysLeuIleGluAlaPheHisProGluArgLeuGluValIleAsnGlu
ATGTCGCTCACCGAGACCATCGAAAAGAAGCTGATCGAGGCCTTCCACCCTGAACGGCTCGAGGTCATCAACGAG
    4098      4108      4118      4128      4138      4148      4158
SerHisGlnHisThrGlyHisGlnProGlyPheAspGlyThrGlyGluSerHisMetArgValArgIleValSer
AGCCATCAGCATACCGGCCATCAGCCGGGCTTCGATGGTACCGGCGAGTCCCACATGCGGGTGCGTATCGTTTCT
    4173      4183      4193      4203      4213      4223      4233
SerAlaPheAlaGlyMetSerArgValAlaArgHisArgAlaIleAsnAspLeuLeuLysProGluLeuAspAla
AGCGCCTTTGCCGGCATGAGCCGTGTCGCCCCGCCACCGCGCCATCAATGATCTCCTGAAGCCAGAACTCGACGCC
    4248      4258      4268      4278      4288      4298      4308
GlyLeuHisAlaLeuAlaValGluProAlaAlaProGlyGluProThrArgTrp**
GGCCTGCATGCGCTCGCCGTCGAGCCGGCAGCCCCGGCGAGCCGACCCGCTGGTAG
    4323      4333      4343      4353      4363
```

FIG. 40G

NAME = COBU			FIRST RESIDUE = 1		LAST RESIDUE = 338	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	9	2.66	1323.62	3.82
2	LEU	L	36	10.65	4071.03	11.75
3	ILE	I	19	5.62	2148.60	6.20
4	MET	M	9	2.66	1179.36	3.40
5	VAL	V	21	6.21	2080.44	6.00
6	SER	S	13	3.85	1131.42	3.26
7	PRO	P	16	4.73	1552.84	4.48
8	THR	T	17	5.03	1717.81	4.96
9	ALA	A	61	18.05	4333.26	12.50
10	TYR	Y	3	0.89	489.19	1.41
11	*	*	0	0.00	0.00	0.00
12	HIS	H	8	2.37	1096.47	3.16
13	GLN	Q	6	1.78	768.35	2.22
14	ASN	N	9	2.66	1026.39	2.96
15	LYS	K	12	3.55	1537.14	4.43
16	ASP	D	13	3.85	1495.35	4.31
17	GLU	E	22	6.51	2838.94	8.19
18	CYS	C	5	1.48	515.05	1.49
19	TRP	W	3	0.89	558.24	1.61
20	ARG	R	16	4.73	2497.62	7.21
21	GLY	G	40	11.83	2280.86	6.58
22	-	-	0	0.00	0.00	0.00

RESIDUES = 338
 MOLECULAR WEIGHT (MONOISOTOPIC) = 34659.9844
 MOLECULAR WEIGHT (AVERAGE) = 34681.9609
 INDEX OF POLARITY (%) = 34.32
 ISOELECTRIC POINT = 6.21
 OD 260 (1mg/ml) = 0.416 OD 280 (1mg/ml) = 0.584

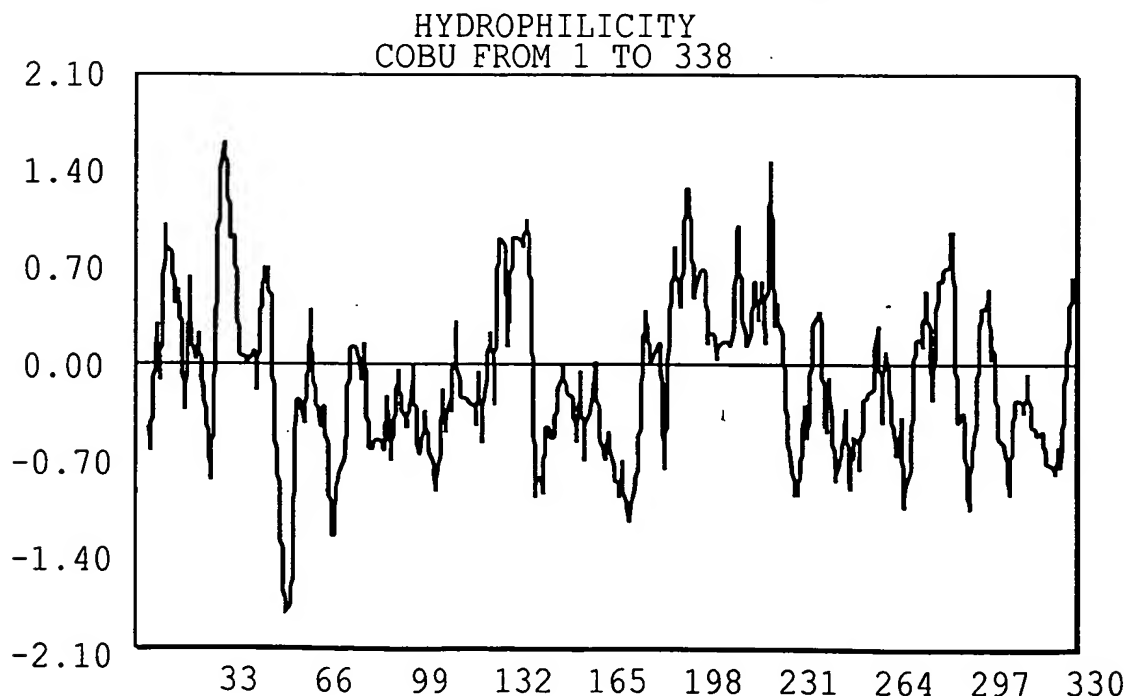


FIG. 41A

cobU GENE (SEQ ID NO: 37) AND COBU PROTEIN (SEQ ID NO: 38)
 SEQUENCE OF THE 3855-BP SstI-SstI-BamHI FRAGMENT
 FROM 2099 TO 3115

```

MetSerAlaSerGlyLeuProPheAspAspPheArgGluLeuLeuArgAsnLeuProGlyProAspAlaAlaAla
ATGAGTGCCAGCGGCTGCCGTTTGATGATTTTCGCGAATTGTTGCGCAACCTGCCGGGCCCCGGATGCCGCAGCC
  2108      2118      2128      2138      2148      2158      2168
LeuValAlaAlaArgGluArgAspAlaGlnLeuThrLysProProGlyAlaLeuGlyArgLeuGluGluIleAla
CTCGTTGCCGCGGGAGCGGGACGCCCAGCTGACGAAGCCGCCGGGCGCGCTCGGCCGCCTCGAGGAAATCGCC
  2183      2193      2203      2213      2223      2233      2243
PheTrpLeuAlaAlaTrpThrGlyLysAlaProValValAsnArgProLeuValAlaIlePheAlaGlyAsnHis
TTCTGGCTCGCCGCTGGACGGGCAAGGCGCCGGTGGTCAACCGCCGCTGGTGGCGATCTTTGCCGGCAACCAC
  2258      2268      2278      2288      2298      2308      2318
GlyValThrArgGlnGlyValThrProPheProSerSerValThrAlaGlnMetValGluAsnPheAlaAlaGly
GGCGTCACCCGCCAGGGGTGACCCGTTCCCGTCATCCGTCACCGCACAGATGGTCGAGAATTTTGCCGCCGGT
  2333      2343      2353      2363      2373      2383      2393
GlyAlaAlaIleAsnGlnIleCysValSerHisAspLeuGlyLeuLysValPheAspLeuAlaLeuGluTyrPro
GGCGTGCATCAACCAGATCTGCGTCAGCCACGACCTCGGGCTGAAGGTCTTCGACCTCGCACTCGAATACCCG
  2408      2418      2428      2438      2448      2458      2468
ThrGlyAspIleThrGluGluAlaAlaLeuSerGluArgAspCysAlaAlaThrMetAlaPheGlyMetGluAla
ACCGGTGATATCACCGAGGAAGCCGCGCTGTCCGAGCGGATTGCGCCGCGACCATGGCCTTTGGCATGGAGGCG
  2483      2493      2503      2513      2523      2533      2543
IleAlaGlyGlyThrAspLeuLeuCysIleGlyGluMetGlyIleGlyAsnThrThrIleAlaAlaAlaIleAsn
ATTGCCGGCGGCACGGATCTTCTGTGCATCGGCGAAATGGGCATCGGCAACACCACGATCGCGGCCGCGATCAAT
  2558      2568      2578      2588      2598      2608      2618
LeuGlyLeuTyrGlyGlyThrAlaGluGluTrpValGlyProGlyThrGlySerGluGlyGluValLeuLysArg
CTCGGCCTTTATGGTGGCACGGCCGAAGAATGGGTCCGTTCCGGGTACCGGCTCCGAGGGCGAGGTGCTGAAGCGC
  2633      2643      2653      2663      2673      2683      2693
LysIleAlaAlaValGluLysAlaValAlaLeuHisArgAspHisLeuSerAspProLeuGluLeuMetArgArg
AAGATCGCCGCGGTGCAAAAGGCCGTGGCGCTGCATCGCGATCACCTGTCCGATCCGCTCGAACTGATGCGTCGC
  2708      2718      2728      2738      2748      2758      2768
LeuGlyGlyArgGluIleAlaAlaMetAlaGlyAlaIleLeuAlaAlaArgValGlnLysValProValIleIle
CTCGGCGGTCTGAGATCGCGGCCATGGCTGGCGCCATCCTGGCCGCCCCGCTCCAGAAGGTACCTGTCATCATC
  2783      2793      2803      2813      2823      2833      2843
AspGlyTyrValAlaThrAlaAlaAlaSerIleLeuLysAlaAlaAsnProSerAlaLeuAspHisCysLeuIle
GACGGCTACGTGGCGACCGCTGCGGCTTCGATCCTGAAGGCGGCCAACCCGTCGGCCCTCGACCATTGCCTGATC
  2858      2868      2878      2888      2898      2908      2918
GlyHisValSerGlyGluProGlyHisLeuArgAlaIleGluLysLeuGlyLysThrProLeuLeuAlaLeuGly
GGCATGTTTCGGGCGAACCGGGGCATCTGCGCGGATCGAGAAGCTCGGCAAGACGCCGCTGCTGGCACTCGGC
  2933      2943      2953      2963      2973      2983      2993
MetArgLeuGlyGluGlyThrGlyAlaAlaLeuAlaAlaGlyIleValLysAlaAlaAlaAlaCysHisSerGly
ATGCGGCTTGCGAAGGCACGGGCGCGGCCCTTGCCGCCGATCGTCAAGGCGGCGGCCGCTTGCCACAGCGGC
  3008      3018      3028      3038      3048      3058      3068
MetAlaThrPheAlaGlnAlaGlyValSerAsnLysGlu***
ATGGCGACCTTTGCCAGGCGCGGTCAGCAACAAGGAATAG
  3083      3093      3103      3113

```

FIG. 41B

NAME = COBV			FIRST RESIDUE = 1		LAST RESIDUE = 302	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	18	5.96	2647.23	8.64
2	LEU	L	39	12.91	4410.28	14.39
3	ILE	I	13	4.30	1470.09	4.80
4	MET	M	10	3.31	1310.41	4.28
5	VAL	V	23	7.62	2278.57	7.44
6	SER	S	18	5.96	1566.58	5.11
7	PRO	P	12	3.97	1164.63	3.80
8	THR	T	10	3.31	1010.48	3.30
9	ALA	A	63	20.86	4475.34	14.61
10	TYR	Y	3	0.99	489.19	1.60
11	*	*	0	0.00	0.00	0.00
12	HIS	H	3	0.99	411.18	1.34
13	GLN	Q	6	1.99	768.35	2.51
14	ASN	N	2	0.66	228.09	0.74
15	LYS	K	5	1.66	640.47	2.09
16	ASP	D	10	3.31	1150.27	3.75
17	GLU	E	7	2.32	903.30	2.95
18	CYS	C	3	0.99	309.03	1.01
19	TRP	W	2	0.66	372.16	1.21
20	ARG	R	19	6.29	2965.92	9.68
21	GLY	G	36	11.92	2052.77	6.70
22	-	-	0	0.00	0.00	0.00

RESIDUES = 302
 MOLECULAR WEIGHT (MONOISOTOPIC) = 30642.3359
 MOLECULAR WEIGHT (AVERAGE) = 30662.0820
 INDEX OF POLARITY (%) = 26.49
 ISOELECTRIC POINT = 9.97
 OD 260 (1mg/ml) = 0.391 OD 280 (1mg/ml) = 0.479

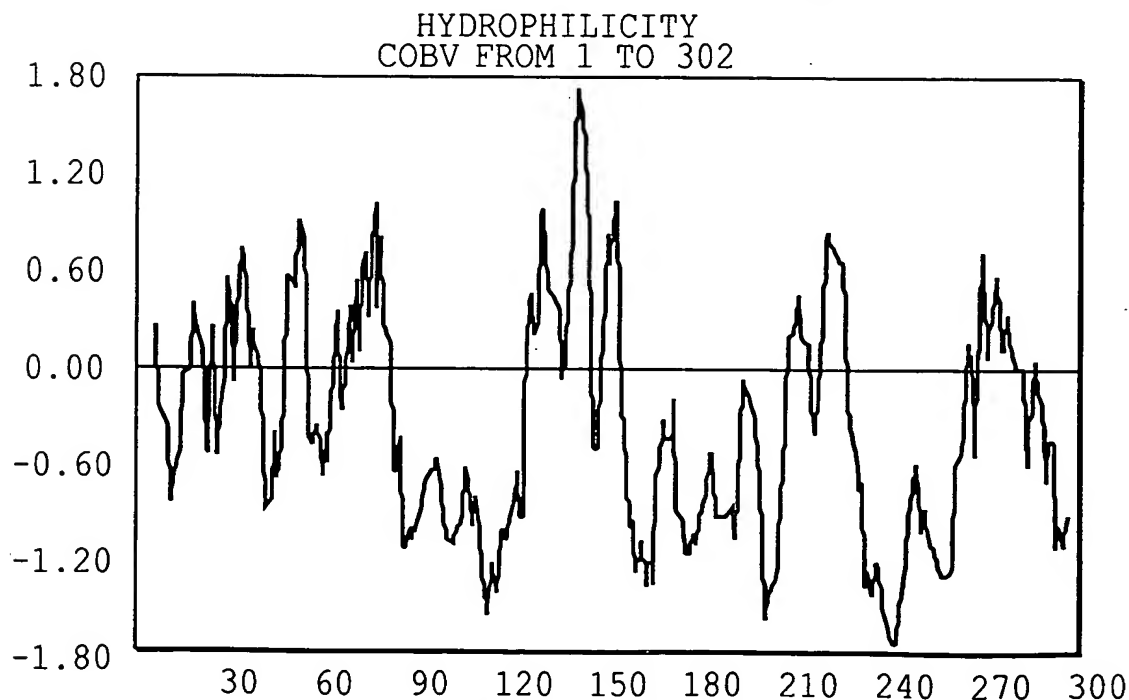


FIG. 41C

cobV GENE (SEQ ID NO: 39) AND COBV PROTEIN (SEQ ID NO: 40)
SEQUENCE OF THE 3855-BP BamHI-SstI-SstI FRAGMENT
FROM 1885 TO 2793

MetLysGlyLeuGlyLysValCysAlaAlaLeuAlaGlyAlaProAlaArgLeuAlaArgLeuAlaAlaValGly
ATGAAAGGTCTGGGAAAAGTGTGCGCGGCCCTCGCGGGTGCGCCCGCCGCTTGCGAGACTCGCCGCAGTCGGC
1894 1904 1914 1924 1934 1944 1954
GluPheGlySerGlyArgValAsnGlyArgAsnLeuGlyGlyAlaMetGlyPheValGlyAspPheCysAspAsp
GAGTTTGGCAGCGGCCGGGTCAATGGCCGCAACCTTGAGGAGCCATGGGCTTCGTTGGCGATTTCTGTGACGAT
1969 1979 1989 1999 2009 2019 2029
ValAlaArgSerIleGlyPheLeuSerArgIleProMetProAlaArgHisPheGluGlyTyrAspGlyArgLeu
GTGGCGCGTTCGATCGGCTTTCTCAGCCGCATCCCGATGCCTGCGCGTCATTTTCAAGGTTATGACGGGCGCTTG
2044 2054 2064 2074 2084 2094 2104
SerArgAlaValArgAlaPheProPheAlaGlyLeuAlaIleAlaLeuProSerAlaAlaValAlaMetAlaLeu
AGCCGTGCCGTGCGCGCCTTCCCGTTCGCCGGCCTGGCGATCGCACTGCCATCGCGGCCGTCGCCATGGCCCTC
2119 2129 2139 2149 2159 2169 2179
MetAlaLeuGlnValSerSerLeuPheAlaAlaPheValValValAlaIleGlnAlaLeuValThrGlyAlaLeu
ATGGCGCTGCAGGTGAGCTCGCTCTTTGCCGCTTTGTCTGTCGTCGCCATCCAGGCGCTTGTACCGGCGCACTG
2194 2204 2214 2224 2234 2244 2254
HisGluAspGlyLeuGlyAspThrAlaAspGlyPheGlyGlyGlyArgAspArgGluAlaAlaLeuAlaIleMet
CACGAAGACGGGCTTGGCGACACGGCCGATGGCTTTGGCGGCGGTGCGGACCGCGAAGCGGCGCTTGCATCATG
2269 2279 2289 2299 2309 2319 2329
LysAspSerArgIleGlyThrTyrAlaAlaValAlaLeuIleLeuSerPheGlyLeuArgValSerAlaPheAla
AAGGACAGCCGCATCGGCACCTATGCGGCCGTCGCCCTCATCTCTCTCGGCCTGCGCGTCTCGGCGTTGCC
2344 2354 2364 2374 2384 2394 2404
SerIleLeuProLeuPheSerProLeuGlyAlaAlaMetAlaIleLeuGlyAlaAlaCysLeuSerArgAlaAla
TCGATCCTGCCGCTCTTTTCGCCGCTGGGCGCCGCCATGGCGATCCTTGGCGCAGCGTGCCTCAGCCGCGCCGCC
2419 2429 2439 2449 2459 2469 2479
MetValTrpHisTrpSerSerLeuProProAlaArgSerSerGlyValAlaAlaSerAlaGlyGluProGluPro
ATGGTCTGGCACTGGTCGTCGCTTCCGCCGGCGCGCAGCAGCGGCGTCGCGGCTCGGCTGGCGAGCCGGAACCG
2494 2504 2514 2524 2534 2544 2554
AlaAlaThrArgPheAlaLeuAlaPheGlyLeuLeuValAlaMetLeuLeuPheTyrLeuAlaGlnValProAla
GCGGCCACGCGCTTCGCGCTTGCCTTCGGAATTCTTGTCGCAATGCTGCTCTTCTATCTCGCGCAGGTCCCGGCC
2569 2579 2589 2599 2609 2619 2629
LeuGlyValIleAlaAlaLeuValAlaPheLeuAlaThrValLysGlyPheAlaArgLeuAlaMetArgLysIle
CTCGGCGTCATCGCGGCGCTCGTCGATTCCTTGCCACGGTCAAGGGCTTCGCTCGGCTGGCGATGCGCAAGATC
2644 2654 2664 2674 2684 2694 2704
GlyGlyGlnThrGlyAspThrIleGlyAlaThrGlnGlnLeuThrGluIleAlaValLeuGlyAlaLeuAlaLeu
GGCGGACAAACGGGCGACAGATCGGGGCGACGAGCAACTGACCGAAATCGCGTGCTCGGTGCCCTTGCCTG
2719 2729 2739 2749 2759 2769 2779
ThrVal***
ACGGTTTGA

FIG. 41D

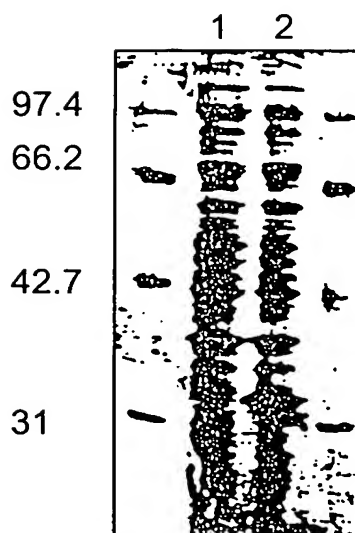


FIG. 42A

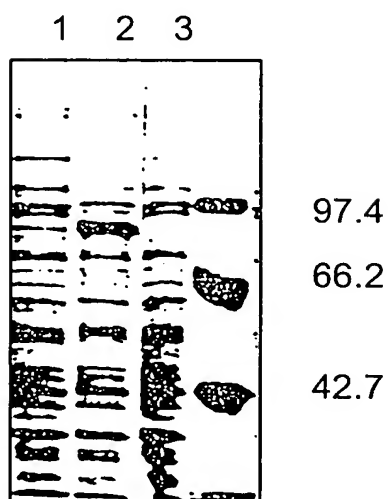


FIG. 42B

SEQUENCE		LENGTH = 13144		FROM 1 TO 13144	
10	20	30	40	50	60
GAGCTCGAAG	GGGCTTCCGC	CCCGATCGCT	GGCGTTAGCC	GACGTTTCGAC	GTGCGGATGA
CTCGAGCTTC	CCCGAAGGCG	GGGCTAGCGA	CCGCAATCGG	CTGCAAGCTG	CACGCCTACT
70	80	90	100	110	120
CGCCGAGCGG	GCCGAAGGGC	GCGTCGACGA	CGAGGTTGCG	TACGCGCGAC	TGGCTGGACG
GCGGCTCGCC	CGGCTTCCCG	CGCAGCTGCT	GCTCCAACGC	ATGCGCGCTG	ACCGACCTGC
130	140	150	160	170	180
GAACCTTCGA	GTTCCAGGCG	ATCTGAACGA	AATTGGGCTT	GCTGAAAATA	TACAGCATGG
CTTGGAAGCT	CAAGGTCCGC	TAGACTTGCT	TTAACCCGAA	CGACTTTTAT	ATGTCGTACC
190	200	210	220	230	240
ACATGAACCT	TGAGAGGCCG	GAGGCCTATC	CTCCGGGGCG	TGTTGCTATG	CCGCTGATAT
TGTACTTGGA	ACTCTCCGGC	CTCCGGATAG	GAGGCCCCGC	ACAACGATAC	GGCGACTATA
250	260	270	280	290	300
AGGTGTGCGC	TGCAAAAAAT	TGAATGCCAA	ACTCGCCACG	CCATGTCGCA	TTCTGGCTAT
TCCACACGCG	ACGTTTTTTA	ACTTACGGTT	TGAGCGGTGC	GGTACAGCGT	AAGACCGATA
310	320	330	340	350	360
CGGCCGCGAC	ATTTTCGACA	AGCCTTGCGA	AAGCGCGAAA	CAATGCGTGA	AAGGGCTTTG
GCCGGCGCTG	TAAAAGCTGT	TCGGAACGCT	TTCGCGCTTT	GTTACGCACT	TTCCCCGAAAC
370	380	390	400	410	420
TCAATTGCGG	CGAAATCGTG	TCGAAACAGA	CCTTTGCCCG	TGCCCCGTTTC	AGTGTTACCG
AGTTAACGCC	GCTTTAGCAC	AGCTTTGTCT	GGAAACGGCG	ACGGGCAAAG	TCACAATGGC
430	440	450	460	470	480
ATGGCCGCAT	GACACGCAGG	ATCATGTTGC	AGGGAACCGG	CTCGGATGTC	GGAAAATCGG
TACCGGCGTA	CTGTGCGTCC	TAGTACAACG	TCCCTTGGCC	GAGCCTACAG	CCTTTTAGCC
490	500	510	520	530	540
TATTGGTGGC	GGGGCTCTGC	CGGCTTGCCG	CCAATCAGGG	CCTGAAGGTC	CGGCCGTTCA
ATAACCACCG	CCCCGAGACG	GCCGAACGGC	GGTTAGTCCC	GGAATTCCAG	GCCGGCAAGT
550	560	570	580	590	600
AGCCGCAGAA	CATGTCGAAC	AACGCCGCCG	TTTCCGACGA	CGGCGGCGAC	ATCGGCCCGC
TCGGCGTCTT	GTACAGCTTG	TTGCGGCGGC	AAAGGCTGCT	GCCGCCGCTC	TAGCCGGCGC

FIG. 43A

610	620	630	640	650	660
CGCAATGGCT	GCAGGCGCTG	GCCGCGCGCG	TGCCGTCGTC	GGTGCACATG	AACCCGGTGC
GCGTTACCGA	CGTCCGCGAC	CGGCGCGCGC	ACGGCAGCAG	CCACGTGTAC	TTGGGCCACG
670	680	690	700	710	720
TCCTGAAGCC	GCAGTCGGAC	GTGGGCAGCC	AGATCGTCGT	TCAGGGCAAG	GTCGCCGGGC
AGGACTTCGG	CGTCAGCCTG	CACCCGTCGG	TCTAGCAGCA	AGTCCCGTTC	CAGCGGCCCG
730	740	750	760	770	780
AGGCCAGGGG	GCGGGAATAT	CAGGCGCTCA	AGCCCAAGCT	GCTGGGCGCC	GTCATGGAGA
TCCGGTCCCC	CGCCCTTATA	GTCCGCGAGT	TCGGGTTCGA	CGACCCGCGG	CAGTACCTCT
790	800	810	820	830	840
GTTTCGAACA	AATATCGGCC	GGTGCCGATC	TCGTGGTGGT	CGAAGGCGCC	GGCTCGCCGG
CAAAGCTTGT	TTATAGCCGG	CCACGGCTAG	AGCACCACCA	GCTTCCGCGG	CCGAGCGGCC
850	860	870	880	890	900
CCGAAATCAA	CCTCAGGCCC	GCGGACATCG	CCAATATGGG	CTTTGCGACA	CGGGCCAATG
GGCTTTAGTT	GGAGTCCGGG	CCGCTGTAGC	GGTTATACCC	GAAACGCTGT	GCCCGGTTAC
910	920	930	940	950	960
TGCCGGTCGT	GCTGGTCGGC	GACATCGACC	GCGGGGGGGT	GATCGCCTCG	CTGGTCGGCA
ACGGCCAGCA	CGACCAGCCG	CTGTAGCTGG	CGCCCCCCCA	CTAGCGGAGC	GACCAGCCGT
970	980	990	1000	1010	1020
CGCATGCGAT	CCTGCCCCGAG	GAAGACCGGC	GCATGGTGAC	CGGCTATCTC	ATCAACAAGT
GCGTACGCTA	GGACGGGCTC	CTTCTGGCCG	CGTACCACTG	GCCGATAGAG	TAGTTGTTCA
1030	1040	1050	1060	1070	1080
TCCGCGGGCGA	CGTCACGCTG	TTCGACGACG	GCATTGCTGC	CGTCAACCGC	TACACCGGCT
AGGCGCCGCT	GCAGTGCGAC	AAGCTGCTGC	CGTAACGACG	GCAGTTGGCG	ATGTGGCCGA
1090	1100	1110	1120	1130	1140
GGCCCTGCTT	CGGCGTCGTG	CCGTGGCTGA	AGGCGGCGGC	ACGCCTGCCG	GCGGAAGATT
CCGGGACGAA	GCCGCAGCAC	GGCACC GACT	TCCGCCGCCG	TGCGGACGGC	CGCCTTCTAA
1150	1160	1170	1180	1190	1200
CCGTCGTGCT	GGAGAAGCTG	ACGCGCGGCG	AGGGGCGGGC	GCTGAAGGTT	GCCGTCCCCG
GGCAGCACGA	CCTCTTCGAC	TGCGCGCCCG	TCCCCGCCCG	CGACTTCCAA	CGGCAGGGCC

FIG. 43B

1210	1220	1230	1240	1250	1260
TACTGTCGCG	CATCGCCAAT	TTCGACGACC	TCGATCCGCT	CGCCGCCGAA	CCGGAGATTG
ATGACAGCGC	GTAGCGGTTA	AAGCTGCTGG	AGCTAGGCGA	GCGGCGGCTT	GGCCTCTAAC
1270	1280	1290	1300	1310	1320
ATCTCGTCTT	CGTGCGGCCT	GGCAGTCCCA	TTCCGGTCGA	CGCTGGCCTC	GTCGTCAATC
TAGAGCAGAA	GCACGCCGGA	CCGTCAGGGT	AAGGCCAGCT	GCGACCGGAG	CAGCAGTAAG
1330	1340	1350	1360	1370	1380
CCGGGTCGAA	ATCGACCATC	GGCGACCTCA	TCGATTTCCG	TGCGCAAGGG	TGGGACCGTG
GGCCCAGCTT	TAGCTGGTAG	CCGCTGGAGT	AGCTAAAGGC	ACGCGTTCCC	ACCTTGGCAC
1390	1400	1410	1420	1430	1440
ACCTCGAACG	TCATGTGCGC	CGGGGCGGCC	GGGTCATCGG	CATCTGCGGC	GGCTACCAGA
TGGAGCTTGC	AGTACACGCG	GCCCCGCCGG	CCCAGTAGCC	GTAGACGCCG	CCGATGGTCT
1450	1460	1470	1480	1490	1500
TGCTCGGCCG	GCGCGTCACC	GATCCGCTCG	GCATCGAGGG	CGGCGAACGT	GCGGTCGAGG
ACGAGCCGGC	CGCGCAGTGG	CTAGGCGAGC	CGTAGCTCCC	GCCGCTTGCA	CGCCAGCTCC
1510	1520	1530	1540	1550	1560
GCCTCGGGCT	GCTCGAGGTC	GAGACCGAGA	TGGCGCCGGA	AAAGACGGTG	CGCAACAGCC
CGGAGCCCCA	CGAGCTCCAG	CTCTGGCTCT	ACCGCGGCCT	TTTCTGCCAC	GCGTTGTCGG
1570	1580	1590	1600	1610	1620
GCGCCTGGTC	GCTGGAGCAT	GATGTGGTGC	TCGAAGGCTA	CGAAATCCAT	CTTGGCAAGA
CGCGGACCAG	CGACCTCGTA	CTACACCACG	AGCTTCCGAT	GCTTTAGGTA	GAACCGTTCT
1630	1640	1650	1660	1670	1680
CGCAAGGTGC	GGA CTGTGGC	CGGCCGTCGG	TGCGCATCGA	CAATCGCGCC	GACGGCGCCC
GCGTTCCACG	CCTGACACCG	GCCGGCAGCC	ACGCGTAGCT	GTTAGCGCGG	CTGCCGCGGG
1690	1700	1710	1720	1730	1740
TTTCGGCCGA	TGGCCGCGTG	ATGGGCACCT	ACCTGCATGG	GCTCTTCACC	AGCGACGCCT
AAAGCCGGCT	ACCGGCGCAC	TACCCGTGGA	TGGACGTACC	CGAGAAGTGG	TCGCTGCGGA
1750	1760	1770	1780	1790	1800
ATCGCGGCGC	GCTGCTCAAG	AGTTTCGGCA	TCGAAGGCGG	CGCCAACAAC	TACCGCCAAT
TAGCGCCGCG	CGACGAGTTC	TCAAAGCCGT	AGCTTCCGCC	GCGGTTGTTG	ATGGCGGTTA

FIG. 43C

1810	1820	1830	1840	1850	1860
CGGTCGATGC	GGCGCTCGAC	GATGTCGCGA	ACGAACTGGA	GGCTGTGCTC	GATCGTCGCT
GCCAGCTACG	CCGCGAGCTG	CTACAGCGCT	TGCTTGACCT	CCGACACGAG	CTAGCAGCGA
1870	1880	1890	1900	1910	1920
GGCTGGACGA	GTTGCTCAGG	CACTAGGGAC	GCGGCAACGG	TCAGCCAGCA	GGTCCGGTAC
CCGACCTGCT	CAACGAGTCC	GTGATCCCTG	CGCCGTTGCC	AGTCGGTCGT	CCAGGCCATG
1930	1940	1950	1960	1970	1980
GTCGGGCCCCA	ACAGGAGCAA	CGAGCTTATC	CGACGGAAC	ACGCTGCGAC	ATCGTGCTCC
CAGCCCGGGT	TGTCCTCGTT	GCTCGAATAG	GCTGCCTTGA	TGCGACGCTG	TAGCACGAGG
1990	2000	2010	2020	2030	2040
TCGCTTGCGG	CTTCCCAGAC	TTCCCGCGCG	GCATCCAGGT	TCATCAGGGC	AATCCCCAGG
AGCGAACGCC	GAAGGGTCTG	AAGGGCGCGC	CGTAGGTCCA	AGTAGTCCCG	TTAGGGGTCC
2050	2060	2070	2080	2090	2100
CCGACGATCA	GGTCCGGCCA	GGCCGACTGC	CACAGATAGG	CTGTCGCCAG	ACCCGCGGCG
GGCTGCTAGT	CCAGGCCGGT	CCGGCTGACG	GTGTCTATCC	GACAGCGGTC	TGGGCGCCCG
2110	2120	2130	2140	2150	2160
ATGATGGCCA	CATTGGCGAA	GGCATCGTTG	CGGGCCGAGA	GAAATGCTGC	CCGCGTGAGC
TACTACCGGT	GTAACCGCTT	CCGTAGCAAC	GCCCGGCTCT	CTTTACGACG	GGCGCACTCG
2170	2180	2190	2200	2210	2220
GTGCCGCTCG	TGTGACGGTA	GGCGACGAGC	AGATAGGCGC	AGAAGAGGTT	GACCACCAGC
CACGGCGAGC	AACTGCCAT	CCGCTGCTCG	TCTATCCGCG	TCTTCTCCAA	CTGGTGCTCG
2230	2240	2250	2260	2270	2280
GCCCCAAGTC	CGGTCAGGGA	CAGGGCAAAG	GGCTCTGGCG	GGACCGGATC	CATGAACTTC
CGGGGTTCAG	GCCAGTCCCT	GTCCCGTTTC	CCGAGACCGC	CCTGGCCTAG	GTAATTGAAG
2290	2300	2310	2320	2330	2340
GCCCAGGCCG	TCCAAAGGAA	GGCCAGCGCC	GGTACCAGTA	GAATGAACGC	CATCGCCATG
CGGGTCCGGC	AGGTTTCCTT	CCGGTCGCGG	CCATGGTCAT	CTTACTTGCG	GTAGCGGTAC
2350	2360	2370	2380	2390	2400
CCGACCCGCG	CGCGGGTTCG	CGCCGTCCAG	GCCAGAGCAA	AGAAAATCAG	CATGTTGACG
GGCTGGGCGC	GCGCCCAAGC	GCGGCAGGTC	CGGTCTCGTT	TCTTTTAGTC	GTACAACTGG

FIG. 43D

2410	2420	2430	2440	2450	2460
GAGGCGTCTT	CGAGGAAGTC	GACGCTGTCTG	GCCATGAGGG	ACACCGAGCC	GATCGAAAGC
CTCCGCAGAA	GCTCCTTCAG	CTGCGACAGC	CGGTACTCCC	TGTGGCTCGG	CTAGCTTTTCG
2470	2480	2490	2500	2510	2520
GCGACAAGGA	GTTCGACCCC	GAAATAGCCA	AGGTTCAACA	GGGAGACGAT	GAGGACGACG
CGCTGTTCCT	CAAGCTGGGG	CTTTATCGGT	TCCAAGTTGT	CCCTCTGCTA	CTCCTGCTGC
2530	2540	2550	2560	2570	2580
CGGCGCAGGT	CGGTATCCAC	TCGAAAGGTT	CCCTTTCTGG	CGAGATTCGC	CCTCGGCACT
GCCGCGTCCA	GCCATAGGTG	AGCTTTCCAA	GGGAAAGACC	GCTCTAAGCG	GGAGCCGTGA
2590	2600	2610	2620	2630	2640
TTTTTTGGCG	AGATTCGCCC	TCGGCACTTT	GGCACAGGTG	TTAGCAGCAG	TTTGCTATCC
AAAAAACCGC	TCTAAGCGGG	AGCCGTGAAA	CCGTGTCCAC	AATCGTCGTC	AAACGATAGG
2650	2660	2670	2680	2690	2700
ATAGCACTAG	GTTTCGACAT	CGGTTCCGTT	CACACTGCCG	TCGTGCCTGA	CGCCCGACAA
TATCGTGATC	CAAAGCTGTA	GCCAAGGCAA	GTGTGACGGC	AGCACGGACT	GCGGGCTGTT
2710	2720	2730	2740	2750	2760
ATCGTCGCGT	GGCGCAACTC	GGCCGGGGAG	GCGTCGCATG	CGTCGATTGA	CTTTGGGCTG
TAGCAGCGCA	CCGCGTTGAG	CCGGCCCCCTC	CGCAGCGTAC	GCAGCTAACT	GAAACCCGAC
2770	2780	2790	2800	2810	2820
CCCGCTTCCT	AATCATCAGG	TGTTGGATGG	TTCCCCCTTG	TCGTGGCGAT	CTGGGGGAAT
GGGCGAAGGA	TTAGTAGTCC	ACAACCTACC	AAGGGGGAAC	AGCACCGCTA	GACCCCCTTA
2830	2840	2850	2860	2870	2880
AATTGGGAAT	GTGACGGATG	GACCCAAATC	GGGCATCCTT	ATCGCAGCCG	ACCCCGCGAC
TTAACCCCTA	CACTGCCTAC	CTGGGTTTAG	CCCGTAGGAA	TAGCGTCGGC	TGGGGCGCTG
2890	2900	2910	2920	2930	2940
TGTAGAACGG	TCAGGGTTTCG	CCATCGGGAT	TGGTGCCGGG	CTGTGCGCCG	GTTGCATGGG
ACATCTTGCC	AGTCCCAAGC	GGTAGCCCTA	ACCACGGCCC	GACAGCCGGC	CAACGTACCC
2950	2960	2970	2980	2990	3000
CAATCGGGGC	AGGTCGGGGA	TCAAGCCGGA	AAAGCCACTG	GCGTGGCATC	GTGATCAGCC
GTTAGCCCCG	TCCAGCCCCT	AGTTTCGGCCT	TTTCGGTGAC	CGCACCGTAG	CACTAGTCGG

FIG. 43E

3010	3020	3030	3040	3050	3060
GGGTTTGGAC	GCCTCTTCTT	CTACGAATCG	TCCGCCTTTC	ACGATGTCCC	TCACAGCGCC
CCCAAACCTG	CGGAGAAGAA	GATGCTTAGC	AGGCGGAAAG	TGCTACAGGG	AGTGTCTGCGG
3070	3080	3090	3100	3110	3120
CATGCGTCGG	AGACGACGCG	CAAAGGTTTCG	CTGTGGCACC	GGAAAGACGC	CGGGAAGGTG
GTACGCAGCC	TCTGCTGCGC	GTTTCCAAGC	GACACCGTGG	CCTTTCTGCG	GCCCTTCCAC
3130	3140	3150	3160	3170	3180
AGGCGGGCCG	CTCGGGCCCT	GACATCGGAA	CCTTGCCGTT	TAAGGGCGAG	GCGATGTTTCG
TCCGCCCGGC	GAGCCCGGGA	CTGTAGCCTT	GGAACGGCAA	ATTCCCGCTC	CGCTACAAGC
3190	3200	3210	3220	3230	3240
GCCCGTGACG	CCGTGAGCCA	GGAGACCTGC	CATCCGGCAT	GGGCATTCCG	CCCGAGGGGA
CGGGCACTGC	GGCACTCGGT	CCTCTGGACG	GTAGGCCGTA	CCCGTAAGGC	GGGCTCCCTT
3250	3260	3270	3280	3290	3300
CTTTTGTCTC	CAACGCCATC	ACGGAGGTTG	TTTTGGCTCG	CAGATGTTTT	CAAGAACGCG
GAAAACAGAG	GTTGCGGTAG	TGCCTCCAAC	AAAACCGAGC	GTGTACAAAA	GTTCTTGCGC
3310	3320	3330	3340	3350	3360
CCCGTGGCGC	GTCCGATGGC	TTTTGCCACC	GACGGCTGAT	TTGGGAATGT	TGAGGCAGCC
GGGCACCGCG	CAGGCTACCG	AAAACGGTGG	CTGCCGACTA	AACCCTTACA	ACTCCGTCGG
3370	3380	3390	3400	3410	3420
ACGATGAGCA	GTCTCAGCGC	CGGGCCCGTG	CTGGTCCTTG	GCGGCGCCCCG	TTCCGGCAAG
TGCTACTCGT	CAGAGTCGCG	GCCCGGGCAC	GACCAGGAAC	CGCCGCGGGC	AAGGCCGTTC
3430	3440	3450	3460	3470	3480
TCCAGCTTTT	CCGAGAGGCT	CGTCGAAGCG	TCCGGCTTCA	CCATGCATTA	TGTCGCCACG
AGGTCGAAAA	GGCTCTCCGA	GCAGCTTCGC	AGGCCGAAGT	GGTACGTAAT	ACAGCGGTGC
3490	3500	3510	3520	3530	3540
GGCCGCGCCT	GGGACGACGA	AATGCGCGAG	CGCATCGACC	ATCACCGGAC	GCGCCGCGGC
CCGGCGCGGA	CCCTGCTGCT	TTACGCGCTC	GCGTAGCTGG	TAGTGGCCTG	CGCGGCGCCG
3550	3560	3570	3580	3590	3600
GAGGGCTGGA	CGACGCATGA	GGAGCCGCTC	GATCTCGTTCG	GCATCCTCAG	ACGCATCGAT
CTCCCGACCT	GCTGCGTACT	CCTCGGCGAG	CTAGAGCAGC	CGTAGGAGTC	TGCGTAGCTA

FIG. 43F

3610	3620	3630	3640	3650	3660
GATCCCAGCC	ATGTGGTCCT	GATCGACTGC	CTGACGCTAT	GGGTCACCAA	TCTCATGCTG
CTAGGGTCGG	TACACCAGGA	CTAGCTGACG	GACTGCGATA	CCCAGTGGTT	AGAGTACGAC
3670	3680	3690	3700	3710	3720
GAAGAGCGCG	ACATGACGGC	GGAGTTCGCC	GCCCTTGTTG	CGTATCTGCC	CGAGGCGCGG
CTTCTCGCGC	TGTACTGCCG	CCTCAAGCGG	CGGGAACAAC	GCATAGACGG	GCTCCGCGCC
3730	3740	3750	3760	3770	3780
GCGCGCCTCG	TCTTTGTTTC	CAATGAGGTC	GGCCTCGGCA	TCGTGCCCCA	GAACCGCATG
CGCGCGGAGC	AGAAACAAAG	GTTACTCCAG	CCGGAGCCGT	AGCACGGGCT	CTTGGCGTAC
3790	3800	3810	3820	3830	3840
GCCCCGCGAGT	TTCGCGACCA	TGCCGGCCGG	CTTCACCAGA	TCGTTGCGGA	GAAATCCGCT
CGGGCGCTCA	AAGCGCTGGT	ACGGCCGGCC	GAAGTGGTCT	AGCAACGCCT	CTTTAGGCGA
3850	3860	3870	3880	3890	3900
GAAGTTTACT	TTGTCGCGGC	CGGTTTGCCG	CTGAAAATGA	AGGGTTGATC	CATGACCACT
CTTCAAATGA	AACAGCGCCG	GCCAAACGGC	GACTTTTACT	TCCCAACTAG	GTACTGGTGA
3910	3920	3930	3940	3950	3960
GCGAGAGCCA	ACCAGGGCAA	GATCCC GGCG	ACCGTCATCA	CCGGCTTCCT	CGGCGCCGGC
CGCTCTCGGT	TGGTCCCGTT	CTAGGGCCGC	TGGCAGTAGT	GGCCGAAGGA	GCCGCGGCCG
3970	3980	3990	4000	4010	4020
AAGACGACGA	TGATCCGCAA	CCTGCTGCAG	AACGCCGACG	GCAAGCGCAT	CGGCCTGATC
TTCTGCTGCT	ACTAGGCGTT	GGACGACGTC	TTGCGGCTGC	CGTTCGCGTA	GCCGGACTAG
4030	4040	4050	4060	4070	4080
ATCAACGAGT	TCGGCGATCT	TGGCGTCGAC	GGCGATGTCT	TGAAGGGCTG	CGGTGCCGAC
TAGTTGCTCA	AGCCGCTAGA	ACCGCAGCTG	CCGCTACAGA	ACTTCCCGAC	GCCACGGCTC
4090	4100	4110	4120	4130	4140
GCCTGCACCG	AGGACGACAT	CATCGAGCTC	ACCAATGGCT	GCATCTGCTG	CACCGTGGCT
CGGACGTGGC	TCCTGCTGTA	GTAGCTCGAG	TGTTTACCGA	CGTAGACGAC	GTGGCACCGA
4150	4160	4170	4180	4190	4200
GACGATTTCA	TCCCGACCAT	GACGAAGCTG	CTCGAGCGTG	AAAACCGTCC	TGACCACATC
CTGCTAAAGT	AGGGCTGGTA	CTGCTTCGAC	GAGCTCGCAC	TTTTGGCAGG	ACTGGTGTAG

FIG. 43G

4210	4220	4230	4240	4250	4260
ATCATCGAAA	CCTCGGGCCT	TGCCCTGCCG	CAGCCGCTGA	TCGCCGCTTT	CAACTGGCCG
TAGTAGCTTT	GGAGCCCGGA	ACGGGACGGC	GTCGGCGACT	AGCGGCGAAA	GTTGACCGGC
4270	4280	4290	4300	4310	4320
GATATCCGCA	GCGAAGTGAC	CGTCGATGGC	GTCGTCACCG	TGGTCGACAG	CGCCGCCGTT
CTATAGGCGT	CGCTTCACTG	GCAGCTACCG	CAGCAGTGGC	ACCAGCTGTC	GCGGCGGCAA
4330	4340	4350	4360	4370	4380
GCCGCTGGCC	GCTTTGCCGA	CGACCACGAC	AAGGTCGATG	CGCTGCGCGT	CGAGGACGAC
CGGCGACCGG	CGAAACGGCT	GCTGGTGCTG	TTCCAGCTAC	GCGACGCGCA	GCTCCTGCTG
4390	4400	4410	4420	4430	4440
AATCTCGATC	ACGAAAGCCC	GATCGAGGAG	CTGTTTCGAGG	ATCAACTGAC	GGCTGCCGAT
TTAGAGCTAG	TGCTTTCGGG	CTAGCTCCTC	GACAAGCTCC	TAGTTGACTG	CCGACGGCTA
4450	4460	4470	4480	4490	4500
CTCATCGTTC	TCAACAAGAC	CGATCTGATC	GATGCCTCCG	GCCTCAAGGC	CGTGCGCGAC
GAGTAGCAAG	AGTTGTTCTG	GCTAGACTAG	CTACGGAGGC	CGGAGTTCCG	GCACGCGCTG
4510	4520	4530	4540	4550	4560
GAGGTGTCTT	CGCGCACCAG	CCGCAAGCCC	ACGATGATCG	AGGCGAAAAA	CGGCGAAGTC
CTCCACAGAA	GCGCGTGGTC	GGCGTTCGGG	TGCTACTAGC	TCCGCTTTTT	GCCGCTTCAG
4570	4580	4590	4600	4610	4620
GCCGCTGCCA	TCCTGCTTGG	CCTCGGTGTC	GGCACGGA	GCGATATCGC	CAACCGCAAG
CGGCGACGGT	AGGACGAACC	GGAGCCACAG	CCGTGCCTTT	CGCTATAGCG	GTTGGCGTTC
4630	4640	4650	4660	4670	4680
TCGCATCACG	AGATGGAGCA	CGAGGCAGGT	GAGGAGCACG	ATCACGACGA	GTTGACAGC
AGCGTAGTGC	TCTACCTCGT	GCTCCGTCCA	CTCCTCGTGC	TAGTGCTGCT	CAAGCTGTGC
4690	4700	4710	4720	4730	4740
TTCGTCGTGC	AGCTCGGTTC	GATCGCCGAT	CCGGCCGCCT	TCATCGATCG	CCTGAAGGGC
AAGCAGCAGC	TCGAGCCAAG	CTAGCGGCTA	GGCCGGCGGA	AGTAGCTAGC	GGACTTCCCC
4750	4760	4770	4780	4790	4800
GTAATCGCGG	AGCACGACGT	TCTGCGCCTC	AAGGGTTTTG	CAGACGTGCC	CGGCAAGCCC
CATTAGCGCC	TCGTGCTGCA	AGACGCGGAG	TTCCCAAAC	GTCTGCACGG	GCCGTTCCGG

FIG. 43H

4810 ATGCGCCTCC TACGCGGAGG	4820 TGATCCAGGC ACTAGGTCCG	4830 GGTCGGCGCC CCAGCCGCGG	4840 CGCATCGACC GCGTAGCTGG	4850 AATATTACGA TTATAATGCT	4860 CCGCGCCTGG GGCGCGGACC
4870 GGCGCTGGCG CCGCGACCGC	4880 AAAAGCGCGG TTTTCGCGCC	4890 TACGCGCCTC ATGCGCGGAG	4900 GTCGTCATCG CAGCAGTAGC	4910 GCCTGCACGA CGGACGTGCT	4920 CATGGACGAG GTACCTGCTC
4930 GCGGCGGTGC CGCCGCCACG	4940 GCGCCGCGAT CGCGGCGCTA	4950 CACCGCGCTC GTGGCGCGAG	4960 GTGTAGATCG CACATCTAGC	4970 TTCTTTGAAT AAGAAACTTA	4980 GAAATGATCT CTTTACTAGA
4990 AACGCATTGA TTGCGTAACT	5000 AATGATGCAG TTACTACGTC	5010 TTCCGGATGG AAGGCCTACC	5020 AGAACGCTTT TCTTGCGAAA	5030 TAGCGTTTTT ATCGCAAAAG	5040 GTTCGGAATT CAAGCCTTAA
5050 GCCCCAACGG CGGGGTTGCC	5060 ACAAGACGAA TGTTCTGCTT	5070 TGCATCTGCT ACGTAGACGA	5080 TCTCGCCCAG AGAGCGGGTC	5090 AAAGGAACGA TTTCCTTGCT	5100 TCGCCGACGG AGCGGCTGCC
5110 CAACGAGGCG GTTGCTCCGC	5120 ATCGACCTTG TAGCTGGAAC	5130 GGCAAACGCC CCGTTTGCGG	5140 GGCCGATATC CCGGCTATAG	5150 CTTTTCCTAT GAAAAGGATA	5160 CGGCCGCCGA GCCGGCGGCT
5170 CACCGAGCTC GTGGCTCGAG	5180 TCCTCGATCG AGGAGCTAGC	5190 CCGCGGCTCA GGCGCCGAGT	5200 CGGCCGACGC GCCGGCTGCG	5210 GACGGAGGCT CTGCCTCCGA	5220 TGAGCCTGCG ACTCGGACGC
5230 CATCGCCAGC GTAGCGGTGC	5240 CTGATGAGCC GACTACTCGG	5250 TGATGCACCC ACTACGTGGG	5260 GATGTCGGTC CTACAGCCAG	5270 GACACTTACG CTGTGAATGC	5280 TCGAGCGCAC AGCTCGCGTG
5290 GGCGCGTCAC CCGCGCAGTG	5300 GCCAAGCTGA CGGTTGACT	5310 TCGTCGTCCG AGCAGCAGGC	5320 GCCGCTCGGT CGGCGAGCCA	5330 GGCGCCAGCT CCGCGGTCTG	5340 ATTTCCGTTA TAAAGGCAAT
5350 TCTGCTGGAA AGACGACCTT	5360 GCCCTGCATG CGGGACGTAC	5370 CGGCTGCCGT GCCGACGGCA	5380 CACCCATCGT GTGGGTAGCA	5390 TTCGAGATCG AAGCTCTAGC	5400 CGGTTCTGCC GCCAAGACGG

FIG. 43I

5410	5420	5430	5440	5450	5460
GGGTGACGAC	AAGCCGGATC	CGGGGCTGGA	GCCTTTCTCC	ACCGTCGCAG	CCGACGACCG
CCCCTGCTG	TTCGGCCTAG	GCCCCGACCT	CGGAAAGAGG	TGGCAGCGTC	GGCTGCTGGC
5470	5480	5490	5500	5510	5520
CCAGCGCCTT	TGGGCTTACT	TCACCGAAGG	CGGCTCGGAC	AATGCCGGGC	TGTTTCTCGA
GGTCGCGGAA	ACCCGAATGA	AGTGGCTTCC	GCCGAGCCTG	TTACGGCCCC	ACAAAGAGCT
5530	5540	5550	5560	5570	5580
CTATGCGGCC	GCACTGGTCA	CAGGTGCGGA	GAAGCCGCAG	CCGGCAAAGC	CCCTGTTGAA
GATACGCCGG	CGTGACCAGT	GTCCACGCCT	CTTCGGCGTC	GGCCGTTTCG	GGGACAACTT
5590	5600	5610	5620	5630	5640
GGCCGGCATC	TGGTGGCCGG	GTGCTGGTGT	GATCGGCGTC	AGCGAATGGC	AGTCCCTTGT
CCGGCCGTAG	ACCACCGGCC	CACGACCACA	CTAGCCGCAG	TCGCTTACCG	TCAGGGAACA
5650	5660	5670	5680	5690	5700
TCAGGGACGG	ATGGTAGCGA	GGGAGGGATT	CGAACCCCCG	ACGGTCGGGA	TCTGCTTTTA
AGTCCCTGCC	TACCATCGCT	CCCTCCCTAA	GCTTGGGGGC	TGCCAGCCCT	AGACGAAAAT
5710	5720	5730	5740	5750	5760
CCGCGCGCTC	GTGCAGAGTG	GCGAGACACG	GCCTGTGGAG	GCGCTGATCG	ATGCGCTGGA
GGCGCGCGAG	CACGTCTCAC	CGCTCTGTGC	CGGACACCTC	CGCGACTAGC	TACGCGACCT
5770	5780	5790	5800	5810	5820
GGCTGAAGGT	GTGCGGGCAC	TGCCGGTGTT	TGTCTCAAGC	CTCAAGGATG	CCGTTTCCGT
CCGACTTCCA	CACGCCCGTG	ACGGCCACAA	ACAGAGTTCG	GAGTTCCTAC	GGCAAAGGCA
5830	5840	5850	5860	5870	5880
CGGCACGCTG	CAGGCGATTT	TTTCCGAGGC	CGCACCCGAC	GTGGTGATGA	ACGCCACTGG
GCCGTGCGAC	GTCCGCTAAA	AAAGGCTCCG	GCGTGGGCTG	CACCACTACT	TGCGGTGACC
5890	5900	5910	5920	5930	5940
CTTTGCGGTC	TCGTGCCCCG	GTGCCGACCG	TCAGCCGACG	GTGCTGGAAT	CGACCGGTGC
GAAACGCCAG	AGCAGCGGGC	CACGGCTGGC	AGTCGGCTGC	CACGACCTTA	GCTGGCCACG
5950	5960	5970	5980	5990	6000
GCCGGTGCTG	CAGGTGATTT	TCTCCGGCTC	GTCGCGGGCG	CAATGGGAAA	CGTCGCCGCA
CGGCCACGAC	GTCCACTAAA	AGAGGCCGAG	CAGCGCCCGC	GTTACCCTTT	GCAGCGGCGT

FIG. 43J

6010	6020	6030	6040	6050	6060
GGGGCTGATG	GCGCGCGACC	TCGCCATGAA	CGTGGCACTC	CCCGAAGTCG	ATGGCCGCAT
CCCCGACTAC	CGCGCGCTGG	AGCGGTACTT	GCACCGTGAG	GGGCTTCAGC	TACCGGCGTA
6070	6080	6090	6100	6110	6120
CCTTGCGCGC	GCCGTCTCCT	TCAAGGCGGC	GTCGATCTAT	GACGCCAAGG	TGGAGGCCAA
GGAACGCGCG	CGGCAGAGGA	AGTTCCGCCG	CAGCTAGATA	CTGCGGTTCC	ACCTCCGTTT
6130	6140	6150	6160	6170	6180
TATCGTCGGC	CATGAGCCGC	TCGAAGGCCG	GGTGCGCTTT	GCCGCTGATC	TTGCCGTCAA
ATAGCAGCCG	GTACTCGGCG	AGCTTCCGGC	CCACGCGAAA	CGGCGACTAG	AACGGCAGTT
6190	6200	6210	6220	6230	6240
CTGGGCGAAC	GTGCGCCGGG	CAGAGCCCGC	CGAGCGCCGT	ATTGCCATCG	TCATGGCCAA
GACCCGCTTG	CACGCGGCCC	GTCTCGGGCG	GCTCGCGGCA	TAACGGTAGC	AGTACCGGTT
6250	6260	6270	6280	6290	6300
CTATCCGAAC	CGCGACGGTC	GCCTCGGCAA	CGGTGTCGGG	CTCGACACGC	CGGCCGGTAC
GATAGGCTTG	GCGCTGCCAG	CGGAGCCGTT	GCCACAGCCC	GAGCTGTGCG	GCCGGCCATG
6310	6320	6330	6340	6350	6360
CGTCGAGGTG	CTTAGCGCCA	TGGCGCGGGA	AGGCTATGCG	GTCGGTGAGG	TTCCCGCCGA
GCAGCTCCAC	GAATCGCGGT	ACCGCGCCCT	TCCGATACGC	CAGCCACTCC	AAGGGCGGCT
6370	6380	6390	6400	6410	6420
TGGCGACGCG	CTGATCCGCT	TTCTGATGGC	CGGGCCGACC	AATGCGGCGA	GCCATGACCG
ACCGCTGCGC	GACTAGGCGA	AAGACTACCG	GCCC GGCTGG	TTACGCCGCT	CGGTACTGGC
6430	6440	6450	6460	6470	6480
TGAAATCCGC	GAGCGTATTT	CGCTGAACGA	TTACAAAACG	TTCTTCGATT	CGCTTCCGAA
ACTTTAGGCG	CTCGCATAAA	GCGACTTGCT	AATGTTTTGC	AAGAAGCTAA	GCGAAGGCTT
6490	6500	6510	6520	6530	6540
ACAGATAAAG	GATGAAGTTG	CCGGTCGCTG	GGGCGTGCCG	GAGGCCGATC	CCTTTTTCTT
TGTCTATTTT	CTACTTCAAC	GGCCAGCGAC	CCCGCACGGC	CTCCGGCTAG	GGAAAAAGGA
6550	6560	6570	6580	6590	6600
CGATGGCGCC	TTCGCGCTGC	CGCTCGCCCC	CTTCGGCGAG	GTGATCGTCG	GCATCCAACC
GCTACCGCGG	AAGCGCGACG	GCGAGCGGGC	GAAGCCGCTC	CACTAGCAGC	CGTAGGTTGG

FIG. 43K

6610	6620	6630	6640	6650	6660
GGCGCGCGGC	TACAACATCG	ATCCGAAGGA	AAGCTACCAT	TCCCCGGACC	TCGTGCCGCC
CCGCGCGCCG	ATGTTGTAGC	TAGGCTTCCT	TTCGATGGTA	AGGGGCCTGG	AGCACGGCGG
6670	6680	6690	6700	6710	6720
GCATGGCTAT	CTCGCCTTCT	ACGCCCTTCCT	GCGCCAGCAG	TTCGGAGCGC	AGGCGATCGT
CGTACCGATA	GAGCGGAAGA	TGCGGAAGGA	CGCGGTCGTC	AAGCCTCGCG	TCCGCTAGCA
6730	6740	6750	6760	6770	6780
CCACATGGGC	AAGCACGGCA	ATCTCGAATG	GCTGCCGGGC	AAGGCGCTGG	CGCTGTCCGA
GGTGTACCCG	TTCGTGCCGT	TAGAGCTTAC	CGACGGCCCG	TTCCGCGACC	GCGACAGCCT
6790	6800	6810	6820	6830	6840
AACCTGCTAT	CCCGAAGCGA	TCTTCGGGCC	GCTGCCGCAC	ATCTATCCCT	TCATCGTCAA
TTGGACGATA	GGGCTTCGCT	AGAAGCCCGG	CGACGGCGTG	TAGATAGGGA	AGTAGCAGTT
6850	6860	6870	6880	6890	6900
CGATCCGGGC	GAAGGTACGC	AGGCCAAGCG	CCGCACCAGC	GCCGTCATCA	TCGACCACCT
GCTAGGCCCG	CTTCCATGCG	TCCGGTTCGC	GGCGTGGTGC	CGGCAGTAGT	AGCTGGTGG A
6910	6920	6930	6940	6950	6960
GACCCCGCCC	TTGACGCGCG	CCGAATCCTA	CGGCCCGCTC	AAGGATCTGG	AAGCGCTCGT
CTGGGGCGGG	AACTGCGCGC	GGCTTAGGAT	GCCGGGCGAG	TTCCTAGACC	TTGCGGAGCA
6970	6980	6990	7000	7010	7020
CGACGAATAT	TACGACGCCG	CCGGCGGTGA	TCCGCGCCGC	CTCAGGCTGC	TCAGCCGCCA
GCTGCTTATA	ATGCTGCGGC	GGCCGCCACT	AGGCGCGGCG	GAGTCCGACG	AGTCGGCGGT
7030	7040	7050	7060	7070	7080
GATCCTCGAT	CTCGTGCGCG	ACATCGGCCT	CGACAGCGAC	GCAGGCATCG	ACAGGGGCGA
CTAGGAGCTA	GAGCACGCGC	TGTAGCCGGA	GCTGTGCGTG	CGTCCGTAGC	TGTCCCGCT
7090	7100	7110	7120	7130	7140
CAGCGACGAC	AAGGCGCTGG	AAAAGCTCGA	CGCCTATCTC	TGCGACCTCA	AGGAAATGCA
GTCGCTGCTG	TTCCGCGACC	TTTTCGAGCT	GCGGATAGAG	ACGCTGGAGT	TCCTTTACGT
7150	7160	7170	7180	7190	7200
GATCCGCGAC	GGCCTGCACA	TCTTCGGCGT	TGCGCCGGAA	GGGCGGTTGT	TGACGGACCT
CTAGGCGCTG	CCGGACGTGT	AGCCGCCGCA	ACGCGGCCTT	CCCGCCAACA	ACTGCCTGGA

FIG. 43L

7210	7220	7230	7240	7250	7260
CACCGTAGCG	CTGGCGCGCG	TGCCCCGAGG	TCTCGGCGAG	GGCGGCGACC	AGAGCCTGCA
GTGGCATCGC	GACCGCGCGC	ACGGGGCTCC	AGAGCCGCTC	CCGCCGCTGG	TCTCGGACGT
7270	7280	7290	7300	7310	7320
GCGGGCGATC	GCAGCGGATG	CGGGGCTGCG	TGGGTTTGCT	ATTCCCACCT	CGGCGGGGGG
CGCCCGCTAG	CGTCGCCTAC	GCCCCGACGC	ACCCAAACGA	TAAGGGTGGA	GCCGCCCCCC
7330	7340	7350	7360	7370	7380
CAACCCCGCA	CGCGACGCCC	AACCCTTCGA	CCCGCTCGAC	TGCGTCATGT	CCGACACCTG
GTTGGGGCGT	GCGCTGCGGG	TTGGGAAGCT	GGGCGAGCTG	ACGCAGTACA	GGCTGTGGAC
7390	7400	7410	7420	7430	7440
GACAGGCCCC	AAACCGTCCA	TCCTCGCTGA	CCTCTCGGAC	GCCCCCTGGC	GCACCGCCGG
CTGTCCGGGC	TTTGCCAGGT	AGGAGCGACT	GGAGAGCCTG	CGGGGGACCG	CGTGGCGGCC
7450	7460	7470	7480	7490	7500
CGATACGGTC	GAGCGCATCG	AGTTGCTTGC	CGCAAATCTC	GTGTCGGGTG	AACTGGCTTG
GCTATGCCAG	CTCGCGTAGC	TCAACGAACG	GCGTTTAGAG	CACAGCCAC	TTGACCGAAC
7510	7520	7530	7540	7550	7560
CCCGGACCAC	TGGGCCAACA	CCCGCGCCGT	GCTCGGCGAA	ATCGAAACGC	GCCTGAAGCC
GGGCCTGGTG	ACCCGGTTGT	GGGCGCGGCA	CGAGCCGCTT	TAGCTTTGCG	CGGACTTCGG
7570	7580	7590	7600	7610	7620
GTCGATTTC	AACTCGGGTG	CCGCCGAGAT	GACCGGCTTC	CTCACC GGTC	TCAGCGGCCG
CAGCTAAAGT	TTGAGCCAC	GGCGGCTCTA	CTGGCCGAAG	GAGTGGCCAG	AGTCGCCGGC
7630	7640	7650	7660	7670	7680
CTTCGTCGCC	CCCGGTCCAT	CGGGCGCGCC	GACGCGCGGC	CGGCCGGATG	TGTTGCCGAC
GAAGCAGCGG	GGGCCAGGTA	GCCCGCGCGG	CTGCGCGCCG	GCCGGCCTAC	ACAACGGCTG
7690	7700	7710	7720	7730	7740
GGGGCGCAAT	TTCTACTCGG	TCGACAGCCG	CGCCGTGCCG	ACGCCGGCGG	CTTACGAGCT
CCCCGCGTTA	AAGATGAGCC	AGCTGTCGGC	GCGGCACGGC	TGCGGCCGCC	GAATGCTCGA
7750	7760	7770	7780	7790	7800
TGGCAAGAAA	TCGGCCGAGC	TTCTGATCCG	CCGCTACCTG	CAGGACCATG	GCGAATGGCC
ACCGTTCTTT	AGCCGGCTCG	AAGACTAGGC	GGCGATGGAC	GTCCTGGTAC	CGCTTACCGG

FIG. 43M

7810	7820	7830	7840	7850	7860
GTCCTCCTTT	GGCCTGACCG	CCTGGGGCAC	GGCGAACATG	CGCACCGGCG	GCGACGACAT
CAGGAGGAAA	CCGGACTGGC	GGACCCCGTG	CCGCTTGTA	GCGTGCCGCG	CGCTGCTGTA
7870	7880	7890	7900	7910	7920
CGCCCAGGCC	CTGGCGCTGA	TCGGCGCCAA	GCCCACCTGG	GACATGGTCT	CTCGCCGGGT
GCGGGTCCGG	GACCGCGACT	AGCCGCGGTT	CGGGTGGACC	CTGTACCAGA	GAGCGGCCCA
7930	7940	7950	7960	7970	7980
GATGGGCTAC	GAGATCGTGC	CGCTCGCAGT	CCTCGGCCGC	CCACGCGTCG	ACGTGACCTT
CTACCCGATG	CTCTAGCACG	GCGAGCGTCA	GGAGCCGGCG	GGTGCGCAGC	TGCACTGGAA
7990	8000	8010	8020	8030	8040
GCGCATTTCC	GGCTTCTTCC	GCGATGCCTT	CCCGGACCAG	ATCGCGCTCT	TCGACAAGGC
CGCGTAAAGG	CCGAAGAAGG	CGCTACGGAA	GGGCCTGGTC	TAGCGCGAGA	AGCTGTTCCG
8050	8060	8070	8080	8090	8100
GATCCGCGCC	GTCGCGCTGG	AGGAAGACGA	TGCCGACAAC	ATGATCGCCG	CACGCATGCG
CTAGGCGCGG	CAGCGCGACC	TCCTTCTGCT	ACGGCTGTTG	TACTAGCGGC	GTGCGTACGC
8110	8120	8130	8140	8150	8160
GGCGGAAAGC	CGGCGGCTGG	AGGCCGAAGG	CGTGGAAGCC	GCCGAGGCCG	CGCGTCGCGC
CCGCCTTTCG	GCCGCCGACC	TCCGGCTTCC	GCACCTTCGG	CGGCTCCGGC	GCGCAGCGCG
8170	8180	8190	8200	8210	8220
CTCCTACCGC	GTCTTTGGCG	CAAAGCCCGG	TGCCTATGGC	GCCGCCCTGC	AGGCGCTGAT
GAGGATGGCG	CAGAAACCGC	GTTTCGGGCC	ACGGATACCG	CGGCGGGACG	TCCGCGACTA
8230	8240	8250	8260	8270	8280
CGACGAGAAG	GGCTGGGAAA	CCAAAGCAGA	TCTCGCCGAG	GCCTATCTTA	CCTGGGGCGC
GCTGCTCTTC	CCGACCCTTT	GGTTTCGTCT	AGAGCGGCTC	CGGATAGAAT	GGACCCCGCG
8290	8300	8310	8320	8330	8340
CTATGCCTAT	GGCGCCGGCG	AGGAGGGCAA	GGCCGAGCGC	GATCTTTTCG	AGGAGCGCCT
GATACGGATA	CCGCGGCCGC	TCCTCCCGTT	CCGGCTCGCG	CTAGAAAAGC	TCCTCGCGGA
8350	8360	8370	8380	8390	8400
GCGCACGATA	GAGGCGGTGG	TGCAGAACCA	GGACAACCGC	GAGCACGATC	TGCTCGACAG
CGCGTGCTAT	CTCCGCCACC	ACGTCTTGGT	CCTGTTGGCG	CTCGTGCTAG	ACGAGCTGTC

FIG. 43N

8410	8420	8430	8440	8450	8460
CGACGACTAC	TACCAGTTCG	AAGGCGGCAT	GAGCGCTGCC	GCCGAACAGC	TCGGCGGTCA
GCTGCTGATG	ATGGTCAAGC	TTCCGCCGTA	CTCGCGACGG	CGGCTTGTCG	AGCCGCCAGT
8470	8080	8490	8500	8510	8520
CCGTCCGGCG	ATCTACCACA	ACGACCATTTC	CCGTCCGGAA	AAGCCTGTGA	TCCGGTCGCT
GGCAGGCCGC	TAGATGGTGT	TGCTGGTAAG	GGCAGGCCTT	TTCGGACACT	AGGCCAGCGA
8530	8540	8550	8560	8570	8580
CGAAGAAGAG	ATCGGCCGCG	TGGTCCGGGC	CCGCGTCGTC	AATCCCAAGT	GGATCGATGG
GCTTCTTCTC	TAGCCGGCAC	ACCAGGCCCG	GGCGCAGCAG	TTAGGGTTCA	CCTAGCTACC
8590	8600	8610	8620	8630	8640
CGTCATGCGC	CACGGATACA	AGGGCGCCTT	CGAGATCGCT	GCCACGGTCG	ACTACATGTT
GCAGTACGCG	GTGCCTATGT	TCCC GCGGAA	GCTCTAGCGA	CGGTGCCAGC	TGATGTACAA
8650	8660	8670	8680	8690	8700
CGCCTTTGCC	GCGACCACGG	GTGCGGTGCG	CGACCATCAT	TTCGAGGCCG	CTTATCAGGC
GCGGAAACGG	CGCTGGTGCC	CACGCCACGC	GCTGGTAGTA	AAGCTCCGGC	GAATAGTCCG
8710	8720	8730	8740	8750	8760
GTTCAATTGTC	GACGAGCGCG	TGGCTGACTT	CATGCGCGAC	AAGAACCCGG	CCGCCTTTGC
CAAGTAACAG	CTGCTCGCGC	ACCGACTGAA	GTACGCGCTG	TTCTTGGGCC	GGCGGAAACG
8770	8780	8790	8800	8810	8820
CGAGCTTGCC	GAACGCCTGC	TTGAAGCAAT	CGACCGCAAT	CTCTGGACGC	CGCGCTCGAA
GCTCGAACGG	CTTGCGGACG	AACTTCGTTA	GCTGGCGTTA	GAGACCTGCG	GCGCGAGCTT
8830	8840	8850	8860	8870	8880
TTCGGCGCGG	TTTGAACCTG	CCGGCATCGG	CACGGCAGCA	ACCCGGCTTC	GTGCCGGCAA
AAGCCGCGCC	AAACTTGAAC	GGCCGTAGCC	GTGCCGTCGT	TGGGCCGAAG	CACGGCCGTT
8890	8900	8910	8920	8930	8940
TGAATAGAGC	GGTTCCGGGC	TGGCGGTTAT	CCGTCCGGAA	TTGCTTGGA	ACAAAGACCT
ACTTATCTCG	CCAAGGCCCG	ACCGCCAATA	GGCAGGCCTT	AACGAACCTT	TGTTTCTGGA
8950	8960	8970	8980	8990	9000
GGTTCCGTTT	CGCTGCTCAG	TGAAGTGCGA	AAAGGAACCG	AAGCGGGACG	AGGGCGTCTG
CCAAGGCAAA	GCGACGAGTC	ACTTCACGCT	TTTCCTTGGC	TTCACCCTGC	TCCCGCAGAC

FIG. 430

9010	9020	9030	9040	9050	9060
CCCATCCCGA	ACTTGAGAAC	TGAGGGAGTG	ATCATGAGCG	ACGAGACGAC	AGTAGGCGGC
GGGTAGGGCT	TGAACTCTTG	ACTCCCTCAC	TAGTACTCGC	TGCTCTGCTG	TCATCCGCCG
9070	9080	9090	9100	9110	9120
GAAGCCCCGG	CCGAGAAGGA	CGATGCCCGC	CACGCCATGA	AGATGGCGAA	GAAGAAGGCA
CTTCGGGGCC	GGCTCTTCCT	GCTACGGGCG	GTGCGGTACT	TCTACCGCTT	CTTCTTCCGT
9130	9140	9150	9160	9170	9180
GCACGCGAAA	AGATCATGGC	GACGAAGACC	GACGAGAAGG	GTCTGATCAT	CGTCAACACC
CGTGCGCTTT	TCTAGTACCG	CTGCTTCTGG	CTGCTCTTCC	CAGACTAGTA	GCAGTTGTGG
9190	9200	9210	9220	9230	9240
GGCAAAGGCA	AGGGCAAGTC	GACCGCCGGC	TTCGGCATGA	TCTTCCGCCA	TATCGCCCAC
CCGTTTCCGT	TCCCGTTTCC	CTGGCGGGCG	AAGCCGTACT	AGAAGGCGGT	ATAGCGGGTG
9250	9260	9270	9280	9290	9300
GGCATGCCCT	GCGCCGTCGT	GCAGTTCATC	AAGGGTGCGA	TGGCAACCGG	CGAGCGCGAG
CCGTACGGGA	CGCGGCAGCA	CGTCAAGTAG	TTCCCACGCT	ACCGTTGGCC	GCTCGCGCTC
9310	9320	9330	9340	9350	9360
TTGATCGAGA	AGCATTTTCGG	CGATGTCTGC	CAGTTCTACA	CGCTCGGCGA	GGGCTTCACC
AACTAGCTCT	TCGTAAAGCC	GCTACAGACG	GTCAAGATGT	GCGAGCCGCT	CCCGAAGTGG
9370	9380	9390	9400	9410	9420
TGGGAAACGC	AGGATCGCGC	CCGCGATGTT	GCGATGGCTG	AAAAGGCCTG	GGAGAAGGCG
ACCTTTTGCG	TCCTAGCGCG	GGCGCTACAA	CGCTACCGAC	TTTTCCGGAC	CCTCTTCCGC
9430	9440	9450	9460	9470	9480
AAGGAAGTGA	TCCGTGACGA	GCGCAACTCG	ATGGTGCTGC	TCGACGAGAT	CAACATTGCT
TTCCTTGACT	AGGCACTGCT	CGCGTTGAGC	TACCACGACG	AGCTGCTCTA	GTTGTAACGA
9490	9500	9510	9520	9530	9540
CTGCGCTACG	ACTACATCGA	CGTCGCCGAA	GTGGTGCGCT	TCCTGAAGGA	AGAAAAGCCG
GACGCGATGC	TGATGTAGCT	GCAGCGGCTT	CACCACGCGA	AGGACTTCCT	TCTTTTCGGC
9550	9560	9570	9580	9590	9600
CACATGACGC	ATGTGGTGCT	CACCGGCCGC	AACGCGAAAG	AAGACCTGAT	CGAAGTCGCC
GTGTACTGCG	TACACCACGA	GTGGCCGGCG	TTGCGCTTTC	TTCTGGACTA	GCTTCAGCGG

FIG. 43P

9610	9620	9630	9640	9650	9660
GATCTCGTCA	CTGAGATGGA	GCTGATCAAG	CATCCGTTCC	GTTCCGGCAT	CAAGGCGCAG
CTAGAGCAGT	GACTCTACCT	CGACTAGTTC	GTAGGCAAGG	CAAGGCCGTA	GTTCCGCGTC
9670	9680	9690	9700	9710	9720
CAGGGCGTGG	AGTTCTGATG	AGCCAGAGCT	GGCAGTTCTG	GGCGCTGCTT	TCGGCCGCCT
GTCCCGCACC	TCAAGACTAC	TCGGTCTCGA	CCGTCAAGAC	CCGCGACGAA	AGCCGGCGGA
9730	9740	9750	9760	9770	9780
TCGCTGCGCT	CACGGCGGTG	TTTGCCAAGG	TCGGGGTTCG	GCAGATCAAC	TCCGACTTCG
AGCGACGCGA	GTGCCGCCAC	AAACGGTTCC	AGCCCCAACG	CGTCTAGTTG	AGGCTGAAGC
9790	9800	9810	9820	9830	9840
CAACGCTGAT	CCGCACCGTC	GTCATCCTCT	GCGTGATCGC	CGCCATCGTG	GCGGCGACAG
GTTGCGACTA	GGCGTGGCAG	CAGTAGGAGA	CGCACTAGCG	GCGGTAGCAC	CGCCGCTGTC
9850	9860	9870	9880	9890	9900
GGCAGTGGCA	GAAGCCATCG	GAAATCCCGG	GCCGCACCTG	GCTGTTCCCTG	GCGCTGTCAG
CCGTCACCGT	CTTCGGTAGC	CTTTAGGGCC	CGGCGTGGAC	CGACAAGGAC	CGCGACAGTC
9910	9920	9930	9940	9950	9960
GGCTTGCGAC	TGGCGCTTCC	TGGCTTGCTT	ATTTCCGCGC	GCTGAAGCTC	GGCGACGCCG
CCGAACGCTG	ACCGCGAAGG	ACCGAACGGA	TAAAGGCGCG	CGACTTCGAG	CCGCTGCGGC
9970	9980	9990	10000	10010	10020
CCCGCGTGGC	GCCGCTCGAC	AAGCTCTCGA	TCGTCATGGT	CGCGATCTTC	GGCGTGCTCT
GGGCGCACCG	CGGCGAGCTG	TTCGAGAGCT	AGCAGTACCA	GCGCTAGAAG	CCGCACGAGA
10030	10040	10050	10060	10070	10080
TCCTCGGTGA	AAAGCTCAAC	CTGATGAACT	GGCTCGGCGT	CGCCTTCATT	GCCGCCGGGG
AGGAGCCACT	TTTCGAGTTG	GACTACTTGA	CCGAGCCGCA	GCGGAAGTAA	CGGCGGCCCC
10090	10100	10110	10120	10130	10140
CGCTGTTGCT	GGCGGTGTTT	TGAGCGCGCC	TGCTCTGGTG	CCTGTTCACT	GAATGCTCGC
GCGACAACGA	CCGCCACAAA	ACTCGCGCGG	ACGAGACCAC	GGACAAGTGA	CTTACGAGCG
10150	10160	10170	10180	10190	10200
CTCAATCAAT	CCGTAATCCC	GACACATGCA	GTGGTTGTGA	CGAGCGGGAG	GACGGCATGC
GAGTTAGTTA	GGCATTAGGG	CTGTGTACGT	CACCAACACT	GCTCGCCCTC	CTGCCGTACG

FIG. 43Q

10210	10220	10230	10240	10250	10260
AGATTGAAGG	CAATTGGAGC	GAGCGCCTTC	CTGATCCGTC	GGGCCACGTC	GCGCAGTTTCG
TCTAACTTCC	GTTAACCTCG	CTCGCGGAAG	GACTAGGCAG	CCCGGTGCAG	GCGCTCAAGC
10270	10280	10290	10300	10310	10320
GCAGACGCTG	GAAGCGTCGC	AGCCTGAGGG	TGAGCCCTGC	TTCAGACCCA	CCGGCGGACA
CGTCTGCGAC	CTTCGCAGCG	TCGGACTCCC	ACTCGGGACG	AAGTCTGGGT	GGCCGCCTGT
10330	10340	10350	10360	10370	10380
CGCCTGCAAT	AGGCACCGTA	GGCGTCGCCG	AAGACCTTGG	CGAGGTGGGT	TTCCTCCATG
GCGGACGTTA	TCCGTGGCAT	CCGCAGCGGC	TTCTGGAACC	GCTCCACCCA	AAGGAGGTAC
10390	10400	10410	10420	10430	10440
CGGATCTGGT	AGGAAATCGA	GATCCAGGCG	GAGAGCGCCA	GCGCCACCGA	GATGACGTTG
GCCTAGACCA	TCCTTTAGCT	CTAGGTCCGC	CTCTCGCGGT	CGCGGTGGCT	CTACTGCAAC
10450	10460	10470	10480	10490	10500
GGCACCGCCA	TCACCGTGCC	GATCAGCGCG	GTCACCATGC	CGACATAGAT	CGGGTTGCGC
CCGTGGCGGT	AGTGGCACGG	CTAGTCGCGC	CAGTGGTACG	GCTGTATCTA	GCCCAACGCG
10510	10520	10530	10540	10550	10560
GAGAAGGCAT	AGAGGCCTGA	GGTCACAAGC	GGCGCGTCCT	GCTTTTCAGG	GATGCCGATC
CTCTTCCGTA	TCTCCGGACT	CCAGTGTTTCG	CCGCGCAGGA	CGAAAAGTCC	CTACGGCTAG
10570	10580	10590	10600	10610	10620
TTCCAGGAAT	GACGCATCGC	CCATTGCGAC	AGCATCGTCA	GCCCCCGGCC	GAGCGTCATC
AAGGTCCTTA	CTGCGTAGCG	GGTAACGCTG	TCGTAGCAGT	CGGGCGGCGG	CTCGCAGTAG
10630	10640	10650	10660	10670	10680
AGCGCCAGGC	CGACGGCGTG	AAGGATGGGC	GTGTCGAGCG	CCGGGATCCG	GCCGAGGGCA
TCGCGGTCCG	GCTGCCGCAC	TTCTTACCCG	CACAGCTCGC	GGCCCTAGGC	CGGCTCCCGT
10690	10700	10710	10720	10730	10740
GCATCGACGG	AGGCCGGGAG	CATGGCGACC	GCCAGCAGGT	GGATCACCAG	CGCTGCGACG
CGTAGCTGCC	TCCGGCCCTC	GTACCGCTGG	CGGTCGTCCA	CCTAGTGGTCT	GCGACGCTGC
10750	10760	10770	10780	10790	10800
ATCAGGCGGA	AAAGCCTGCC	CGCAAACCCT	TCCGCATCGT	CGCCATAGGT	TAGCACGACC
TAGTCCGCCT	TTTCGGACGG	GCGTTTGGGA	AGGCGTAGCA	GCGGTATCCA	ATCGTGCTGG

FIG. 43R

10810	10820	10830	10840	10850	10860
GGCGAGCGGC	CGGATTGCAC	GCGGCGGAGG	ATCGCCAGCG	CGAGCGTGGA	CAATCCCACG
CCGCTCGCCG	GCCTAACGTG	CGCCGCCTCC	TAGCGGTCGC	GCTCGCACCT	GTTAGGGTGC
10870	10880	10890	10900	10910	10920
ACGAGCATCA	GGATGGTGGG	AAGGGTGGTG	GACATGGAAA	CCTCTGGAGC	GAGCTGACAA
TGCTCGTAGT	CCTACCACCC	TTCCCACCAC	CTGTACCTTT	GGAGACCTCG	CTCGACTGTT
10930	10940	10950	10960	10970	10980
GACAGGAGCG	CACGACGGGT	AGGCGGCCCA	TATGAGCGTC	TACCCGGCGA	AGCATTCTGA
CTGTCCTCGC	GTGCTGCCCA	TCCGCCGGGT	ATACTCGCAG	ATGGGCCGCT	TCGTAAGACT
10990	11000	11010	11020	11030	11040
TCACCTTGCA	ATCTCTAGTA	ACTAGAGGTT	CAAGCGTCGG	ACCTGTCCGA	CTTTCGTCGT
AGTGGAACGT	TAGAGATCAT	TGATCTCCAA	GTTTCGCAGCC	TGGACAGGCT	GAAAGCAGCA
11050	11060	11070	11080	11090	11100
GGTTACCGGA	TCTTATTGCC	AAGCGTTGGA	GGCTGTCATC	GTCGCCCCCG	CCGTGTCGGA
CCAATGGCCT	AGAATAACGG	TTCGCAACCT	CCGACAGTAG	CAGCGGGGGC	GGCACAGCCT
11110	11120	11130	11140	11150	11160
AGGTCGGCAA	AATTCGTCTC	TTGACGGCTG	CTCCTTCCGT	CGAGCGATTG	CATAGGCAGG
TCCAGCCGTT	TTAAGCAGAG	AACTGCCGAC	GAGGAAGGCA	GCTCGCTAAC	GTATCCGTCC
11170	11180	11190	11200	11210	11220
AGGCCGCACC	CATGTTAGAC	CGTCGACAGG	CTAAATACGG	GTGAACCTTG	AAGAATACTC
TCCGGCGTGG	GTACAATCTG	GCAGCTGTCC	GATTTATGCC	CACTTGGAAC	TTCTTATGAG
11230	11240	11250	11260	11270	11280
TCAGAGCTGC	GGTTGGTGTC	GCATCGGTCT	TGCTGTTCTT	GTCATCAGGT	GTGGCGGGGC
AGTCTCGACG	CCAACCACAG	CGTAGCCAGA	ACGACAAGAA	CAGTAGTCCA	CACCGCCCCG
11290	11300	11310	11320	11330	11340
AGGCGCAAAC	CGTGAAGAGC	GGGGCGTCAC	GAGCTCAAGA	AACGACGACC	ACCCAGAAGG
TCCGCGTTTG	GCACTTCTCG	CCCCGCAGTG	CTCGAGTTCT	TTGCTGCTGG	TGGGTCTTCC
11350	11360	11370	11380	11390	11400
CGAAACCGAA	AACTAAAACG	ACGCGCAAGC	AAAGGGCTGC	GGATGAAGCC	AAGGCCAAGG
GCTTTGGCTT	TTGATTTTGC	TGCGCGTTTC	TTTCCCGACG	CCTACTTCGG	TTCCGGTTCC

FIG. 43S

11410	11420	11430	11440	11450	11460
CGCTCGCCGA	AGCGCGCCGT	CCACGGATTT	GCAAGACGCG	GGAGAGCGAA	TGCAGCTATG
GCGAGCGGCT	TCGCGCGGCA	GGTGCCTAAA	CGTTCTGCGC	CCTCTCGCTT	ACGTCGATAAC
11470	11480	11490	11500	11510	11520
GCGCAGGTCC	GGTCGGAGAG	CAGTGCTCGT	GCTGGTCGAA	ATCCGGTGCG	CCTGATCTTG
CGCGTCCAGG	CCAGCCTCTC	GTCACGAGCA	CGACCAGCTT	TAGGCCACGC	GGACTAGAAC
11530	11540	11550	11560	11570	11580
GCATAACTGT	CAGGCGTTGA	CCGCCCCGCG	CCTTCGCGCG	GGCAGGCAAG	CGTGCGTCGC
CGTATTGACA	GTCCGCAACT	GGCGGGCGCT	GGAAGCGCGC	CCGTCCGTTC	GCACGCAGCG
11590	11600	11610	11620	11630	11640
TCGAAGCGAC	GCCTGACGCG	ATAGAAATCA	CGGGTCGCCT	GGTTCGTTCT	GAAAGCTTGG
AGCTTCGCTG	CGGACTGCGC	TATCTTTAGT	GCCCAGCGGA	CCAAGCAAGA	CTTTCGAACC
11650	11660	11670	11680	11690	11700
GATTGGGTTT	AGGTGATGGA	AGCCGGCGTT	GAACGCAAAA	TAATGATCGA	TCTCGAGAAC
CTAACCCAAA	TCCACTACCT	TCGGCCGCAA	CTTGCGTTTT	ATTACTAGCT	AGAGCTCTTG
11710	11720	11730	11740	11750	11760
AGCGCGCTCC	AGTTTGCAAC	CCGAGCACAC	GGCGAACAGA	AGCGTAAGTA	TGACGGTCGG
TCGCGCGAGG	TCAAACGTTG	GGCTCGTG TG	CCGCTTGTCT	TCGCATT CAT	ACTGCCAGCC
11770	11780	11790	11800	11810	11820
CCCTATATCG	TTCATCCGAT	TGCGGTGGCG	GAGATTGTTC	GAAGCGTGCC	CCATACGCC
GGGATATAGC	AAGTAGGCTA	ACGCCACCGC	CTCTAACAAG	CTTCGCACGG	GGTATGCGGG
11830	11840	11850	11860	11870	11880
GAAATGATCG	CCGCAGCGCT	GCTTCACGAT	ACGGTCGAAG	ATACCGACGC	GACGCTGCTG
CTTTACTAGC	GGCGTCGCGA	CGAAGTGCTA	TGCCAGCTTC	TATGGCTGCG	CTGCGACGAC
11890	11900	11910	11920	11930	11940
GAGATCAAGG	AAGCGTTCGG	CCCCAAGGTC	GCAACACTGG	TTGCGTGGCT	CACCGACATA
CTCTAGTTCC	TTCGCAAGCC	GGGGTTCCAG	CGTTGTGACC	AACGCACCGA	GTGGCTGTAT
11950	11960	11970	11980	11990	12000
TCCACTCCGT	TCCACGGCAA	CCGACAGGTG	CGCAAGGAAC	TGGATCGCCA	GCACCTCGCA
AGGTGAGGCA	AGGTGCCGTT	GGCTGTCCAC	GCGTTCCTTG	ACCTAGCGGT	CGTGGAGCGT

FIG. 43T

12010	12020	12030	12040	12050	12060
TCGGCGCCCG	CCGCGGCGAA	AACCGTCAAG	CTCGCCGACC	TGATCGACAA	TGCGATAGCG
AGCCGCGGGC	GGCGCCGCTT	TTGGCAGTTC	GAGCGGCTGG	ACTAGCTGTT	ACGCTATCGC
12070	12080	12090	12100	12110	12120
ATCAAAGCCG	GCGATCCGAA	TTTCTGGAAA	GTGTTTCGGCG	CCGAGATGAA	ACGCTTGCTG
TAGTTTCGGC	CGCTAGGCTT	AAAGACCTTT	CACAAGCCGC	GGCTCTACTT	TGCGAACGAC
12130	12140	12150	12160	12170	12180
GAGGTCTTGG	GCGACGGCGA	CGAGACCCTT	CTCGCAAAGG	CCCGTGCATT	AGCGCCGGAA
CTCCAGAACC	CGCTGCCGCT	GCTCTGGGAA	GAGCGTTTCC	GGGCACGTAA	TCGCGGCCTT
12190	12200	12210	12220	12230	12240
TGAGAGTGCC	GCCGTTTATC	GGCAAGCATG	TCTGTGCCAT	GTCGACCCGG	TCAACCGGTC
ACTCTCACGG	CGGCAAATAG	CCGTTTCGTAC	AGACACGGTA	CAGCTGGGCC	AGTTGGCCAG
12250	12260	12270	12280	12290	12300
ATCCAAGATC	GCAGAACGGA	CATGCATTTG	CGGTTTTGCC	CGCCGGTGTG	GCCCAGCCAC
TAGGTTCTAG	CGTCTTGCCT	GTACGTAAAC	GCCAAAACGG	GCGGCCACAC	CGGGTCGGTG
12310	12320	12330	12340	12350	12360
GCCTCACAGG	CTGCGCGGTT	GCGGCCGTTA	GGACAGCGCA	GAATTTGCCG	ACCGCGCCGC
CGGAGTGTCC	GACGCGCCAA	CGCCGGCAAT	CCTGTGCGCT	CTTAAACGGC	TGGCGCGGCG
12370	12380	12390	12400	12410	12420
GCCTCAATGC	CCCAGCCAGA	TCCGCAAGGG	ATGCGTCGGA	TCTGCGAGCA	GCCGGATCGC
CGGAGTTACG	GGGTCGGTCT	AGGCGTTCCC	TACGCAGCCT	AGACGCTCGT	CGGCCTAGCG
12430	12440	12450	12460	12470	12480
GAGCGCGATC	GAGACGATGA	CGAGCAGCGG	CTTGATGATC	TTGGCGCCCT	TGGCCATGGC
CTCGCGCTAG	CTCTGCTACT	GCTCGTCGCC	GAAC TACTAG	AACCGCGGGA	ACCGGTACCG
12490	12500	12510	12520	12530	12540
ATAGCGCGAG	CCGACCTGGG	CGCCGAGGAA	CTGGCCGAGG	CCCATCAACA	GGCCGACCTT
TATCGCGCTC	GGCTGGACCC	GCGGCTCCTT	GACCGGCTCC	GGGTAGTTGT	CCGGCTGGAA
12550	12560	12570	12580	12590	12600
CCAGAGAACG	GCGCCGAAGA	AGAGGAAGAC	GCCGAAGGCG	CCGACGTTGG	AGCCAAAGTT
GGTCTCTTGC	CGCGGCTTCT	TCTCCTTCTG	CGGCTTCCGC	GGCTGCAACC	TCGGTTTCAA

FIG. 43U

12610	12620	12630	12640	12650	12660
GAGGAAC TTC	GTGTGCGCCG	TCGCCTTCAA	CACGCCGAAG	CCGGCGAGGG	TAACGAAGCC
CTCCTTGAAG	CACACGCGGC	AGCGGAAGTT	GTGCGGCTTC	GGCCGCTCCC	ATTGCTTCGG
12670	12680	12690	12700	12710	12720
GAGCATGAAG	AACGAGCCGG	TGCCGGGGCC	GAAGACGCCG	TCATAAAAGC	CGATTAGCGG
CTCGTACTTC	TTGCTCGGCC	ACGGCCCCGG	CTTCTGCGGC	AGTATTTTCG	GCTAATCGCC
12730	12740	12750	12760	12770	12780
CACCAGTGTC	AGCGTGAAGA	CGAAGGGGGT	GACGCGGCTG	TGCTGGTCGA	CGTCGCCCCAT
GTGGTCACAG	TCGCACTTCT	GCTTCCCCCA	CTGCGCCGAC	ACGACCAGCT	GCAGCGGGTA
12790	12800	12810	12820	12830	12840
GTTCGGCTTC	AGGCCGAAAT	AAAGCGCAAT	GGCGATCAGC	AGAAAGGGCA	GGATCGCCTT
CAAGCCGAAG	TCCGGCTTTA	TTTCGCGTTA	CCGCTAGTCG	TCTTTCCCGT	CCTAGCGGAA
12850	12860	12870	12880	12890	12900
CAGCACGTCG	CCGGGAACGA	TGGTTGCGAG	CAGGGCGCCG	AGCACGGCGC	CGGCGGCCGA
GTCGTGCAGC	GGCCCTTGCT	ACCAACGCTC	GTCCCGCGGC	TCGTGCCGCG	GCCGCCGGCT
12910	12920	12930	12940	12950	12960
CATCAGCGCC	ATCGGCAGCT	GCTCTTTCAG	GTTACAGTGG	CCGCGCCGGG	CATAGGACAG
GTAGTCGCGG	TAGCCGTCGA	CGAGAAAGTC	CAAGTGCACC	GGCGCGGCC	GTATCCTGTC
12970	12980	12990	13000	13010	13020
CGTGGCCGAG	CCGGAGCCGA	ACAATCCCTG	CAGCTTGTTG	GTGCCGAGCG	TCTGCAAGGG
GCACCGGCTC	GGCCTCGGCT	TGTTAGGGAC	GTCGAACAAC	CACGGCTCGC	AGACGTTCCC
13030	13040	13050	13060	13070	13080
CGGGATGCCC	GCAATGAGCA	TGGCCGGAAT	GGTGATCATG	CCACCGCCGC	CGGCGATCGA
GCCCTACGGG	CGTTACTCGT	ACCGGCCTTA	CCACTAGTAC	GGTGGCGGCG	GCCGCTAGCT
13090	13100	13110	13120	13130	13140
ATCGATGAAG	CCTGCGATGA	AGGCGGCGAC	GAACAGGAAG	GCGAGCAGGT	GGAAGGCGAG
TAGCTACTTC	GGACGCTACT	TCCGCCGCTG	CTTGTCCTTC	CGCTCGTCCA	CCTTCCGCTC

ATCT
 TAGA

FIG. 43V

RESTRICTION MAP OF THE 13144-BP SEQUENCE

ApaLI	642,			
EcoRI	8818,			
HindIII	11633,			
MluI	7963,			
NdeI	10950,			
PvuII	12918,			
SfiI	3133,			
SplI	99,			
BglII	8248,	13139,		
KpnI	2315,	6300,		
NotI	5526,	7615,		
SmaI	1322,	9868,		
SspI	4843,	6968,		
XmnI	9313,	12091,		
AatII	1033,	9503,	12773,	
AflIII	550,	7963,	8634,	
BalI	2107,	6236,	12473,	
BamHI	2266,	5416,	10664,	
BspMII	5002,	8494,	8914,	
EcoRV	4263,	4605,	5137,	
NcoI	6318,	7786,	12474,	
NsiI	3467,	5064,	12266,	
PflMI	7870,	10718,	11065,	
XhoI	1512,	4171,	11692,	
ApaI	1928,	3138,	3386,	8551,
AsuII	784,	5670,	8418,	11799,
FspII	784,	5670,	8418,	11799,
MaeI	1883,	2647,	10995,	11002,
NruI	1827,	3794,	10002,	12419,
SauI	852,	7001,	10284,	10517,
BstEII	995,	3642,	8456,	10470, 11041,
Eco47III	6954,	7209,	8434,	10731, 11837,
SacI	5,	4109,	4694,	5169, 11315,
StuI	204,	4081,	8261,	9406, 10515,
BstXI	761,	2982,	3612,	6031, 6232, 9102,
SacII	932,	1025,	2096,	3537, 5184, 12014,
SphI	966,	2740,	5360,	8098, 9246, 10199,

FIG. 44A

RESTRICTION MAP OF THE 13144-BP SEQUENCE

BclI	2992,	4016,	9029,	9164,	9623,	10978,	13053,	
RsaI	101,	1201,	1918,	2313,	4881,	6298,	6856,	
Tth111I	1821,	2424,	4351,	7361,	7904,	12227,	12697,	
PstI	613,	3989,	5832,	5952,	7260,	7782,	8211,	12992,
ClaI	1351,	3596,	4469,	4724,	5748,	6618,	8574,	11687,
	13082,							
FspI	1363,	1551,	1653,	5219,	7841,	7982,	8342,	9760,
	11971,							
HinfI	1137,	2564,	2592,	3025,	5667,	5927,	6467,	6923,
	13079,							
StyI	2488,	3396,	5116,	6105,	6318,	7786,	9745,	10355,
	11389,	11395,	11903,	12468,	12474,			
DdeI	852,	1875,	3373,	3586,	6311,	7001,	7010,	7610,
	8956,	9020,	9611,	10284,	10517,	11220,		
Nsp7524I	554,	966,	2394,	2740,	5360,	7840,	8098,	8638,
	9246,	9553,	10168,	10199,	12210,	12264,		
PvuI	26,	1853,	2453,	4403,	4703,	4728,	5091,	5112,
	5178,	6717,	7269,	9991,	12429,	13077,		
AvaI	975,	1320,	1503,	1512,	3131,	3231,	3709,	3766,
	4171,	4212,	7224,	7573,	9866,	11692,	11720,	
BanII	5,	496,	1723,	1928,	2254,	3138,	3386,	4109,
	4694,	5196,	6207,	6282,	8551,	10296,	11315,	
SalI	83,	1296,	2418,	4045,	4303,	5258,	6959,	7700,
	7967,	8627,	8708,	9198,	11182,	12221,	12766,	
XhoII	2266,	3920,	5416,	5688,	6943,	7020,	7140,	8248,
	10382,	10400,	10664,	11048,	12378,	12398,	13139,	

FIG. 44B

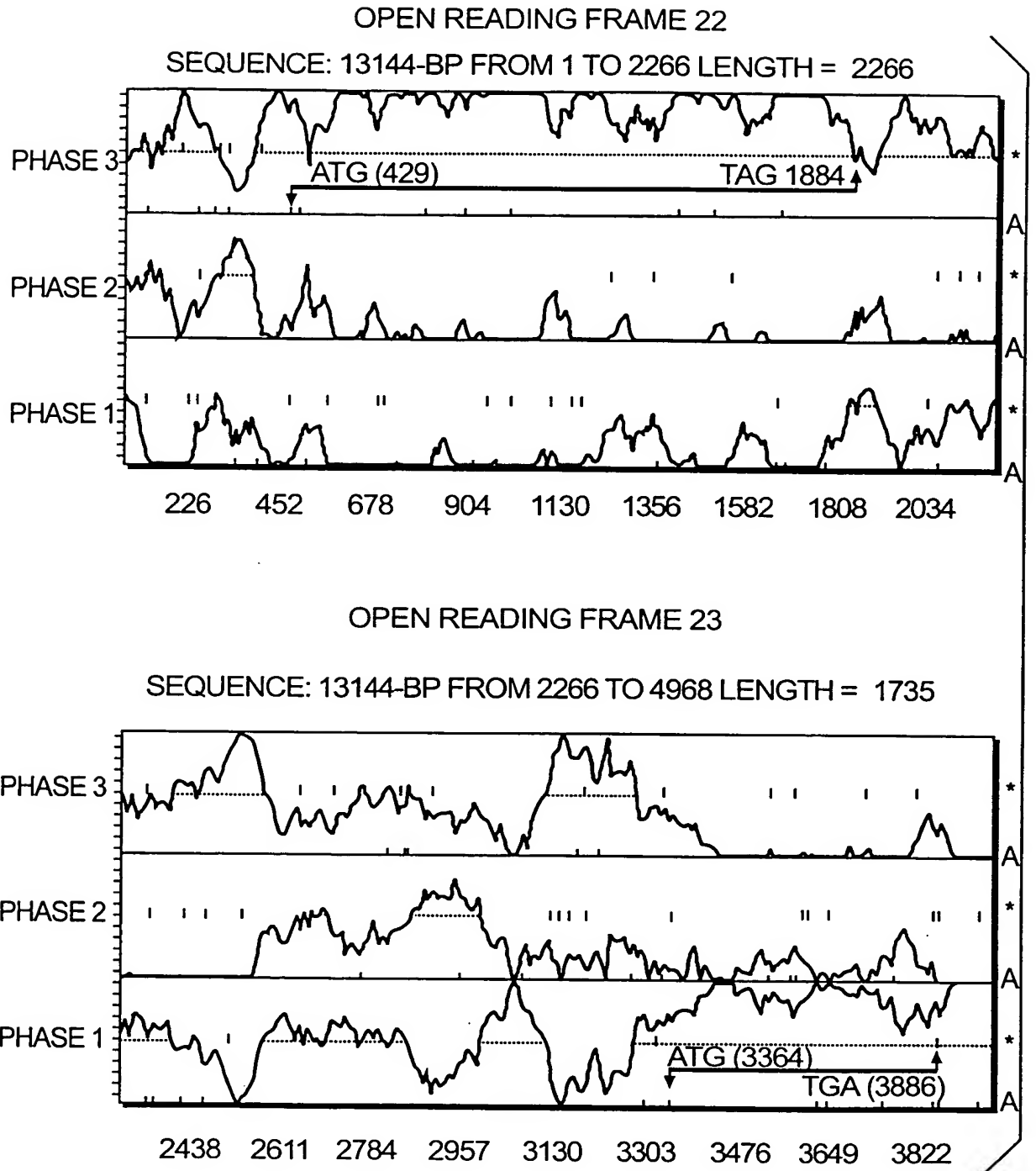


FIG. 45A

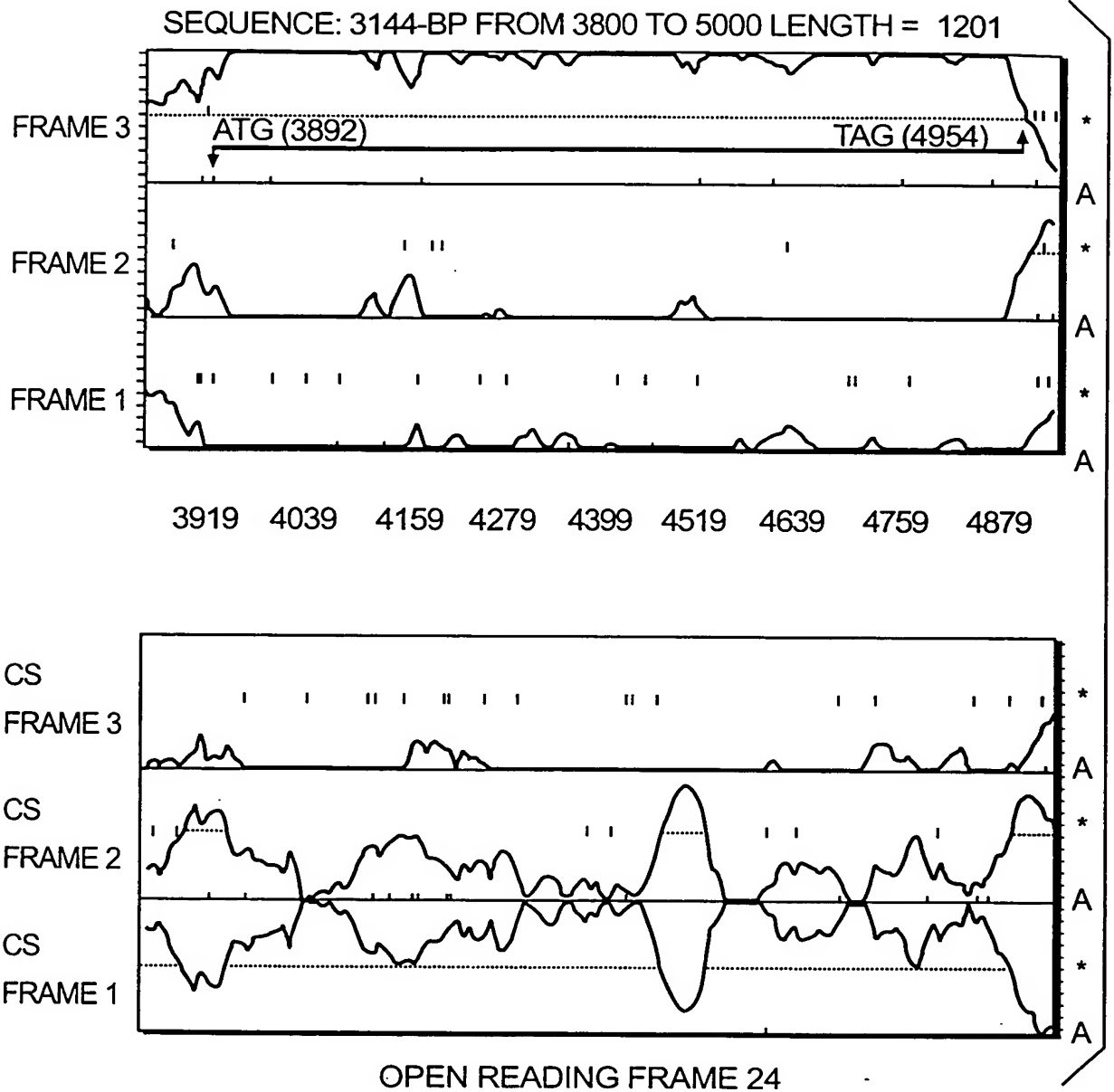


FIG. 45B

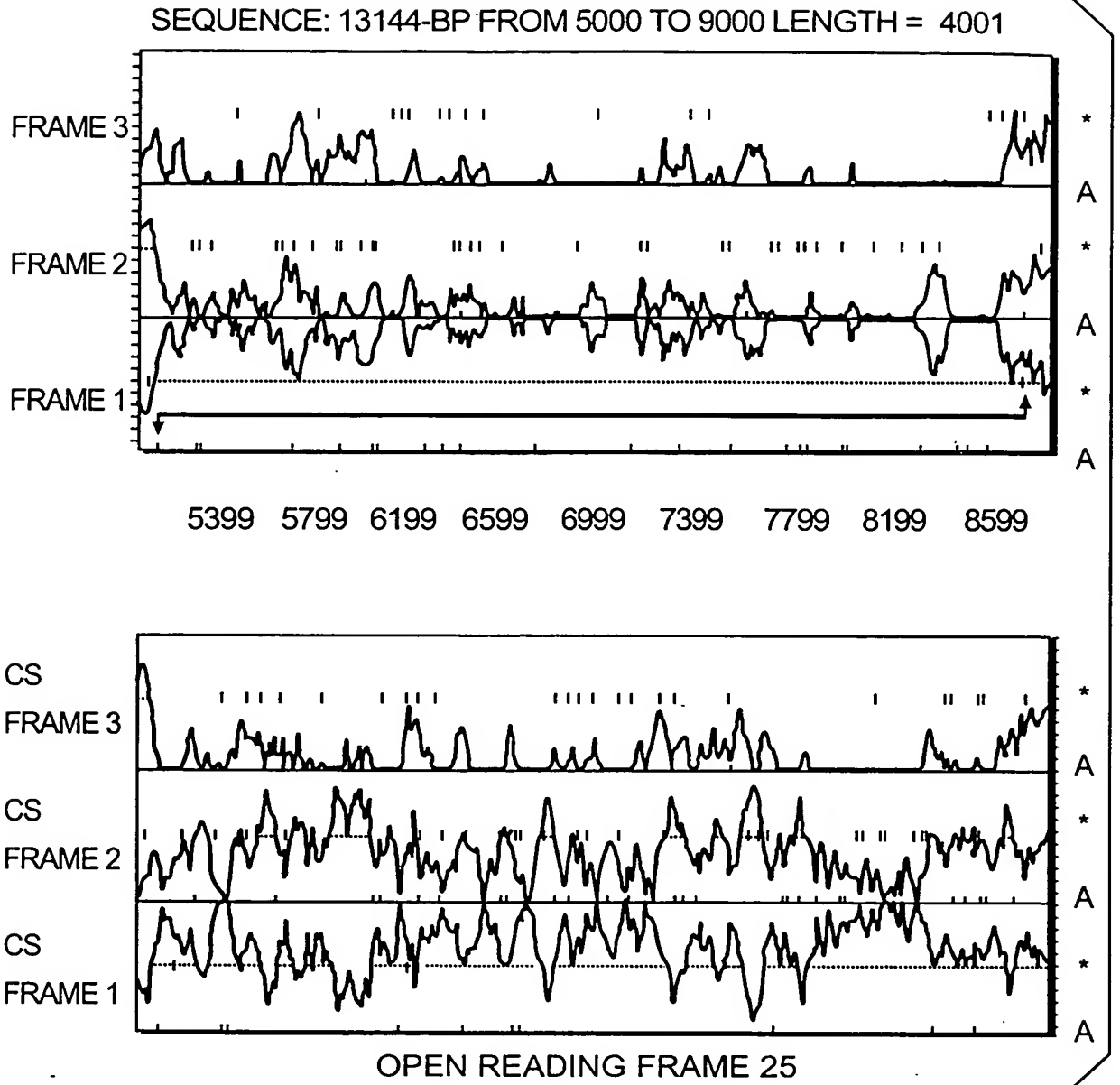


FIG. 45C

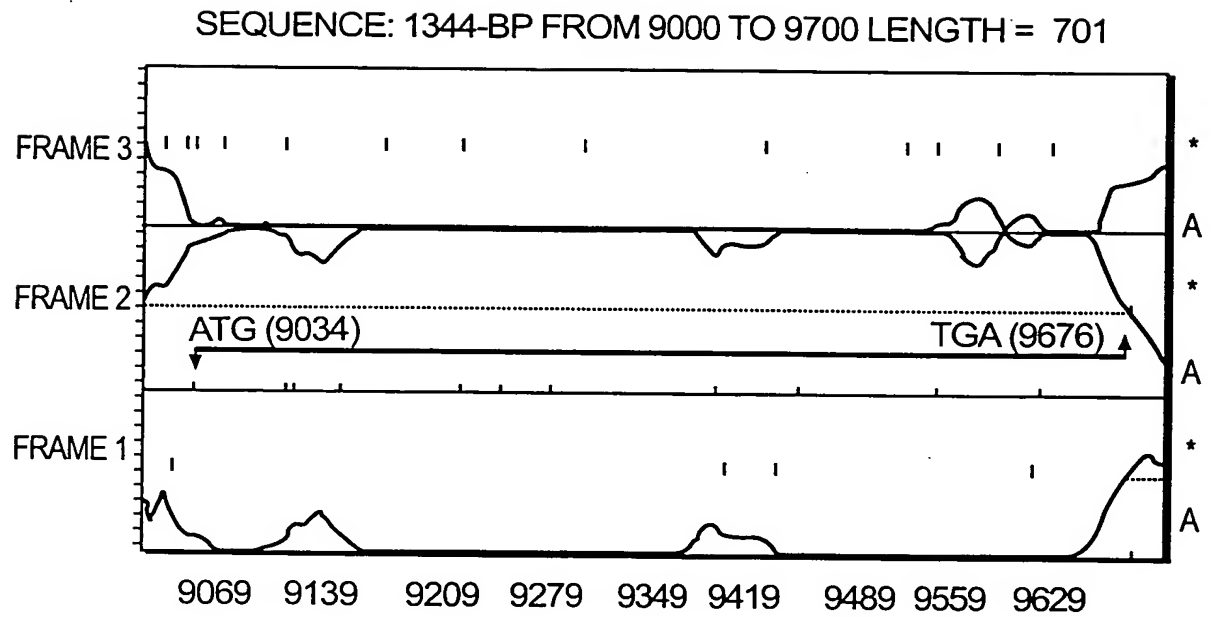


FIG. 45D

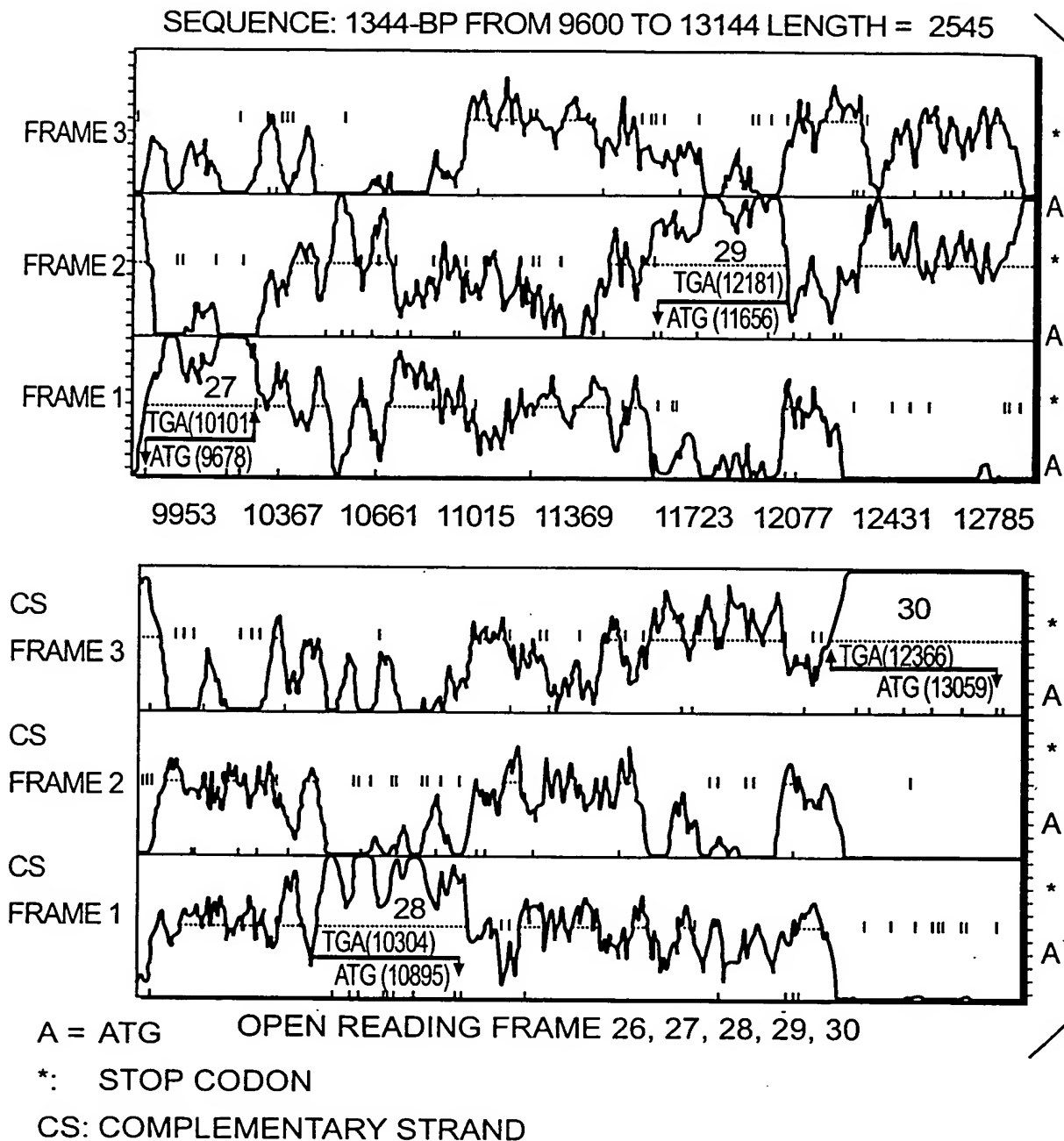


FIG. 45E

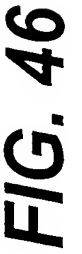


FIG. 46

13144-BP SEQUENCE FROM 429 TO 1328 cobQ GENE

M T R R I M L Q G T G S D V G K S V L V A G L C R
ATGACACGCAGGATCATGTTGCAGGGAACCGGCTCGGATGTCGAAAATCGGTATTGGTGGCGGGGCTCTGCCGG
429 439 449 459 469 479 489 499
L A A N Q G L K V R P F K P Q N M S N N A A V S D
CTTGCCGCCAATCAGGGCCTGAAGGTCCGGCCGTTCAAGCCGAGAACATGTCGAACAACGCCGCCGTTTCCGAC
504 514 524 534 544 554 564 574
D G G E I G R A Q W L Q A L A A R V P S S V H M N
GACGGCGGCGAGATCGGCCGCGCAATGGCTGCAGGCGCTGGCCGCGCGCTGCCGTCGTCGGTGCACATGAAC
579 589 599 609 619 629 639 649
P V L L K P Q S D V G S Q I V V Q G K V A G Q A R
CCGGTGCTCCTGAAGCCGAGTCGGACGTGGGCAGCCAGATCGTCGTTCAAGGCAAGGTGCCCGGGCAGGCCAGG
654 664 674 684 694 704 714 724
G R E Y Q A L K P K L L G A V M E S F E Q I S A G
GGGCGGGAATATCAGGCGCTCAAGCCCAAGCTGCTGGGCGCCGTCATGGAGAGTTTGAACAAATATCGGCCCGGT
729 739 749 759 769 779 789 799
A D L V V V E G A G S P A E I N L R P G D I A N M
GCCGATCTCGTGGTGGTCAAGGCGCCGGCTCGCCGCGCGAAATCAACCTCAGGCCCGGCGACATCGCCAATATG
804 814 824 834 844 854 864 874
G F A T R A N V P V V L V G D I D R G G V I R S L
GGCTTTGCGACACGGGCAATGTGCCGGTCGTGCTGGTCGGCGACATCGACCGCGGGGGGTGATCGCCTCGCTG
879 889 899 909 919 929 939 949
V G T H A I L P E E D R R M V T G Y L I N K F R G
GTCGGCACGCATGCGATCCTGCCCCGAGGAAGACCGGCGCATGGTGACCGGTATCTCATCAACAAGTTCCGCGGC
954 964 974 984 994 1004 1014 1024
D V T L F D D G I A A V N R Y T G W P C F G V V P
GACGTCACGCTGTTGACGACGGCATTGCTGCCGTCAACCGCTACACCGGTGGCCCTGCTTCGGCGTCGTGCCG
1029 1039 1049 1059 1069 1079 1089 1099
W L K A A A R L P A E D S V V L E K L T R G E G R
TGGCTGAAGGCGGCGGCACGCCTGCCGGCGGAAGATTCCGTGCTGCTGGAGAAGCTGACGCGCGGCGAGGGGCGG
1104 1114 1124 1134 1144 1154 1164 1174
A L K V A V P V L S R I A N F D D L D P L A A E P
GCGCTGAAGGTTGCCGTCCCGGTACTGTCGCGCATCGCCAATTTGACGACCTCGATCCGCTCGCCGCCGAACCG
1179 1189 1199 1209 1219 1229 1239 1249
E I D L V F V R P G S P I P V D A G L V V I P G S
GAGATTGATCTCGTCTTCGTGCGGCCTGGCAGTCCCATTCGGTCGACGCTGGCCTCGTCGTATTCCCGGGTCG
1254 1264 1274 1284 1294 1304 1314 1324

FIG. 47A

13144-BP SEQUENCE FROM 1329 TO 1886 cobQ GENE

```

      K S T I G D L I D F R A Q G W D R D L E R H V R R
      AAATCGACCATCGGCGACCTCATCGATTTCCGTGCGCAAGGGTGGGACCGTGACCTCGAACGTCATGTGCGCCGG
1329      1339      1349      1359      1369      1379      1389      1399
      G G R V I G I C G G Y Q M L G R R V T D P L G I E
      GCGGCGCGGGTCATCGGCATCTGCGGCGGCTACCAGATGCTCGGCCGGCGCGTCACCGATCCGCTCGGCATCGAG
1404      1414      1424      1434      1444      1454      1464      1474
      G G E R A V E G L G L L E V E T E M A P E K T V R
      GCGGCGGAACGTGCGGTCGAGGGCCTCGGGCTGCTCGAGGTCGAGACCGAGATGGCGCCGAAAAGACGGTGCGC
1479      1489      1499      1509      1519      1529      1539      1549
      N S R A W S L E H D V V L E G Y E I H L G K T Q G
      AACAGCCGCGCCTGGTCGCTGGAGCATGATGTGGTGCTCGAAGGCTACGAAATCCATCTTGGCAAGACGCAAGGT
1554      1564      1574      1584      1594      1604      1614      1624
      A D C G R P S V R I D N R A D G A L S A D G R V M
      GCGGACTGTGGCCGGCCGTCGGTGCGCATCGACAATCGCGCCGACGGCGCCCTTTCGGCCGATGGCCGCGTGATG
1629      1639      1649      1659      1669      1679      1689      1699
      G T Y L H G L F T S D A Y R G A L L K S F G I E G
      GGCACCTACCTGCATGGGCTCTTACCAGCGACGCCTATCGCGGCGCGCTGCTCAAGAGTTTCGGCATCGAAGGC
1704      1714      1724      1734      1744      1754      1764      1774
      G A N N Y R Q S V D A A L D D V A N E L E A V L D
      GCGCCAACAACACTACCGCCAATCGGTCGATGCGGGCGCTCGACGATGTCGCGAACGAACGAGGCTGTGCTCGAT
1779      1789      1799      1809      1819      1829      1839      1849
      R R W L D E L L R H * (SEQ ID NO:43)
      CGTCGCTGGCTGGACGAGTTGCTCAGGCACTAG (SEQ ID NO:42)
1854      1864      1874      1884
```

FIG. 47B

COBQ PROTEIN

FIRST RESIDUE = 1
 LAST RESIDUE = 485

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	11	2.27	1617.75	3.11
2	LEU	L	50	10.31	5654.20	10.88
3	ILE	I	23	4.74	2600.93	5.01
4	MET	M	10	2.06	1310.41	2.52
5	VAL	V	50	10.31	4953.42	9.53
6	SER	S	24	4.95	2088.77	4.02
7	PRO	P	23	4.74	2232.21	4.30
8	THR	T	15	3.09	1515.72	2.92
9	ALA	A	49	10.10	3480.82	6.70
10	TYR	Y	8	1.65	1304.51	2.51
11	*	*	0	0.00	0.00	0.00
12	HIS	H	7	1.44	959.41	1.85
13	GLN	Q	15	3.09	1920.88	3.70
14	ASN	N	16	3.30	1824.69	3.51
15	LYS	K	15	3.09	1921.42	3.70
16	ASP	D	34	7.01	3910.92	7.53
17	GLU	E	28	5.77	3613.19	6.96
18	CYS	C	4	0.82	412.04	0.79
19	TRP	W	6	1.24	1116.48	2.15
20	ARG	R	40	8.25	6244.04	12.02
21	GLY	G	57	11.75	3250.22	6.26
22	-	-	0	0.00	0.00	0.00

RESIDUES = 485
 MOLECULAR WEIGHT (MONOISOTOPIC) = 51950.1016
 MOLECULAR WEIGHT (AVERAGE) = 51982.3711
 INDEX OF POLARITY (%) = 40.00
 ISOELECTRIC POINT = 6.16
 OD 260 (1mg/ml) = 0.558 OD 280 (1mg/ml) = 0.825

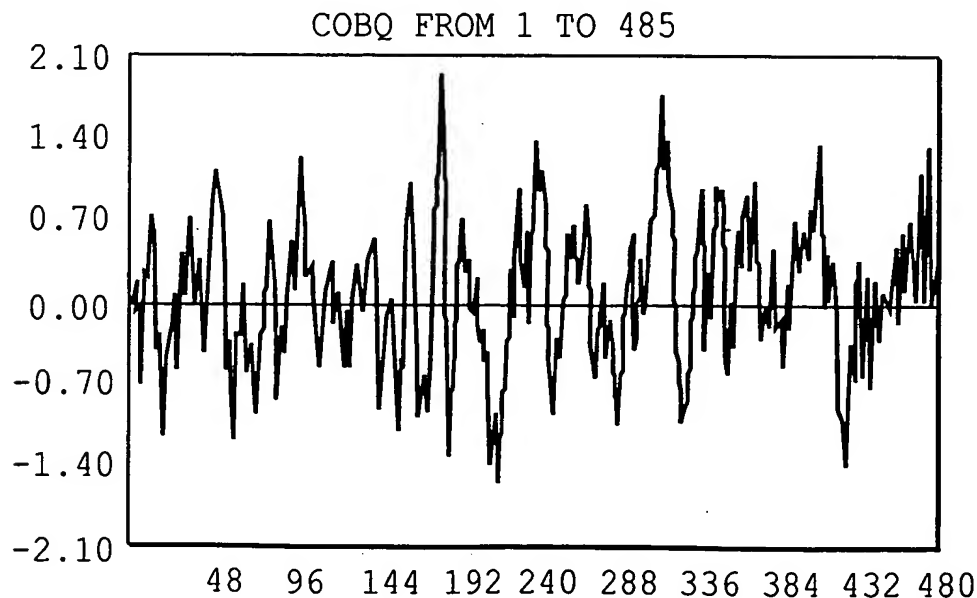


FIG. 47C

13144-BP SEQUENCE FROM 3364 TO 3888 cobP GENE

```

      M S S L S A G P V L V L G G A R S G K S S F S E R
      ATGAGCAGTCTCAGCGCCGGGCGGCTGCTGGTCCTTGGCGGCGCCCGTTCCGGCAAGTCCAGCTTTTCCGAGAGG
3364      3374      3384      3394      3404      3414      3424      3434
      L V E A S G F T M H Y V A T G R A W D D E M R E R
      CTCGTCAAGCGTCCGGCTTCACCATGCATTATGTGCCACGGGCGCGCTGGGACGACGAAATGCGCGAGCGC
3439      3449      3459      3469      3479      3489      3499      3509
      I D H H R T R R G E G W T T H E E P L D L V G I L
      ATCGACCATCACCGGACGCGCGCGGCGGAGGGCTGGACGACGCATGAGGAGCCGCTCGATCTCGTCGGCATCCTC
3514      3524      3534      3544      3554      3564      3574      3584
      R R I D D P S H V V L I D C L T L W V T N L M L E
      AGACGCATCGATGATCCCAGCCATGTGGTCCTGATCGACTGCCTGACGCTATGGGTCACCAATCTCATGCTGGAA
3589      3599      3609      3619      3629      3639      3649      3659
      E R D M T A E F A A L V A Y L P E A R A R L V F V
      GAGCGCGACATGACGGCGGAGTTCGCCGCCCTTGTTGCGTATCTGCCCGAGGCGCGGGCGCGCCTCGTCTTTGTT
3664      3674      3684      3694      3704      3714      3724      3734
      S N E V G L G I V P E N R M A R E F R D H A G R L
      TCCAATGAGGTCGGCCTCGGCATCGTGCCCGAGAACCGCATGGCCCGCGAGTTTCGCGACCATGCCGCGCGGCTT
3739      3749      3759      3769      3779      3789      3799      3809
      H Q I V A E K S A E V Y F V R A G L P L K M K G *
      CACCAGATCGTTGCGGAGAAATCCGCTGAAGTTTACTTTGTGCGGCGCGGTTTGCCGCTGAAAATGAAGGGTTGA
3814      3824      3834      3844      3854      3864      3874      3884

```

(SEQ ID NO: 45)
(SEQ ID NO: 44)

FIG. 47D

COBP PROTEIN FIRST RESIDUE = 1
 LAST RESIDUE = 174

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	6	3.45	882.41	4.54
2	LEU	L	19	10.92	2148.60	11.06
3	ILE	I	6	3.45	678.50	3.49
4	MET	M	7	4.02	917.28	4.72
5	VAL	V	16	9.20	1585.09	8.16
6	SER	S	11	6.32	957.35	4.93
7	PRO	P	6	3.45	582.32	3.00
8	THR	T	8	4.60	808.38	4.16
9	ALA	A	17	9.77	1207.63	6.22
10	TYR	Y	3	1.72	489.19	2.52
11	*	*	0	0.00	0.00	0.00
12	HIS	H	7	4.02	959.41	4.94
13	GLN	Q	1	0.57	128.06	0.66
14	ASN	N	3	1.72	342.13	1.76
15	LYS	K	4	2.30	512.38	2.64
16	ASP	D	9	5.17	1035.24	5.33
17	GLU	E	16	9.20	2064.68	10.63
18	CYS	C	1	0.57	103.01	0.53
19	TRP	W	3	1.72	558.24	2.87
20	ARG	R	17	9.77	2653.72	13.66
21	GLY	G	14	8.05	798.30	4.11
22	-	-	0	0.00	0.00	0.00

RESIDUES = 174
 MOLECULAR WEIGHT (MONOISOTOPIC) = 19429.9473
 MOLECULAR WEIGHT (AVERAGE) = 19442.2637
 INDEX OF POLARITY (%) = 43.68
 ISOELECTRIC POINT = 6.71
 OD 260 (1mg/ml) = 0.720 OD 280 (1mg/ml) = 1.042

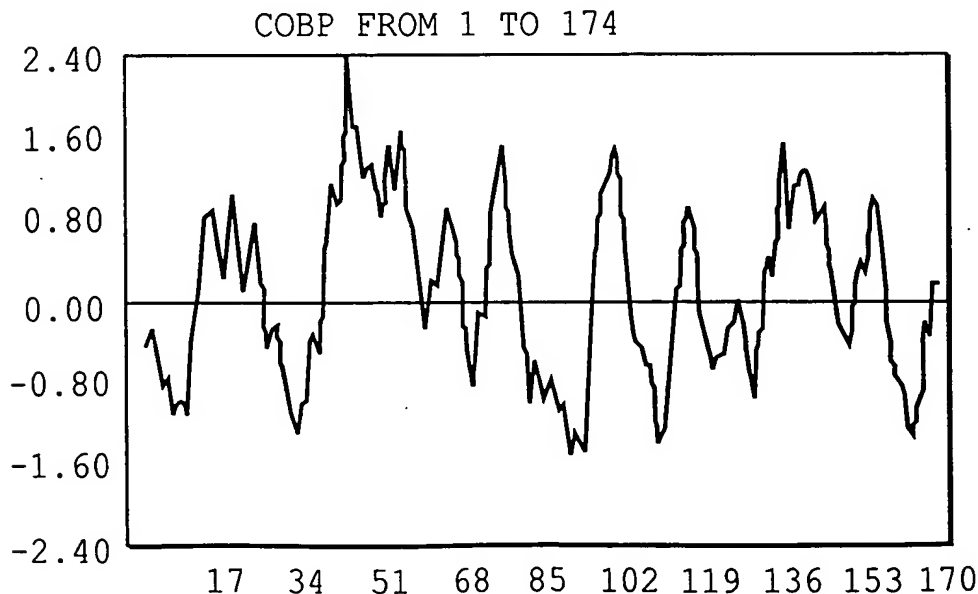


FIG. 47E

13144-BP SEQUENCE FROM 3892 TO 4956 cobW GENE

M T T A R A N Q G K I P A T V I T G F L G A G K T
ATGACCACTGCGAGAGCCAACCAGGGCAAGATCCCGGCGACCGTCATCACCGGCTTCCTCGGCGCCGGCAAGACG
3892 3902 3912 3922 3932 3942 3952 3962
T M I R N L L Q N A D G K R I G L I I N E F G D L
ACGATGATCCGCAACCTGCTGCAGAACGCCGACGGCAAGCGCATCGGCCTGATCATCAACGAGTTCCGGCGATCTT
3967 3977 3987 3997 4007 4017 4027 4037
G V D G D V L K G C G A E A C T E D D I I E L T N
GGCGTCGACGGCGATGTCTTGAAGGGCTGCGGTGCCGAGGCCTGCACCGAGGACGACATCATCGAGCTCACCAAT
4042 4052 4062 4072 4082 4092 4102 4112
G C I C C T V A D D F I P T M T K L L E R E N R P
GGCTGCATCTGCTGCACCGTGGCTGACGATTTTCATCCCGACCATGACGAAGCTGCTCGAGCGTGAAAACCGTCCT
4117 4127 4137 4147 4157 4167 4177 4187
D H I I I E T S G L A L P Q P L I A A F N W P D I
GACCACATCATCATCGAAACCTCGGGCCTTGCCCTGCCGACGCCGCTGATCGCCGCTTTCAACTGGCCGGATATC
4192 4202 4212 4222 4232 4242 4252 4262
R S E V T V D G V V T V V D S A A V A A G R F A D
CGCAGCGAAGTGACCGTCGATGGCGTCGTACCGTGGTCGACAGCGCCGCGTTGCCGCTGGCCGCTTTGCCGAC
4267 4277 4287 4297 4307 4317 4327 4337
D H D K V D A L R V E D D N L D H E S P I E E L F
GACCACGACAAGGTCGATGCGCTGCGCGTCGAGGACGACAATCTCGATCAGAAAGCCCGATCGAGGAGCTGTTC
4342 4352 4362 4372 4382 4392 4402 4412
E D Q L T A A D L I V L N K T D L I D A S G L K A
GAGGATCAACTGACGGCTGCCGATCTCATCGTTCTCAACAAGACCGATCTGATCGATGCCTCCGGCCTCAAGGCC
4417 4427 4437 4447 4457 4467 4477 4487
V R D E V S S R T S R K P T M I E A K N G E V A A
GTGCGCGACGAGGTGTCTTCGCGCACCAGCCGCAAGCCCACGATGATCGAGGCGAAAAACGGCGAAGTCGCCGCT
4492 4502 4512 4522 4532 4542 4552 4562
A I L L G L G V G T E S D I A N R K S H H E M E H
GCCATCCTGCTTGGCCTCGGTGTCGGCACGGAAAGCGATATCGCCAACCGCAAGTCGCATCACGAGATGGAGCAC
4567 4577 4587 4597 4607 4617 4627 4637
E A G E E H D H D E F D S F V V E L G S I A D P A
GAGGCAGGTGAGGAGCAGATCACGACGAGTTCGACAGCTTCGTCGTCGAGCTCGGTTCGATCGCCGATCCGGCC
4642 4652 4662 4672 4682 4692 4702 4712
A F I D R L K G V I A E H D V L R L K G F A D V P
GCCTTCATCGATCGCCTGAAGGGCGTAATCGCGGAGCACGACGTTCTGCGCCTCAAGGGTTTTGCAGACGTGCC
4717 4727 4737 4747 4757 4767 4777 4787

FIG. 47F

G K P M R L L I Q R V G A R I D Q Y Y D R A W G A
GGCAAGCCGATGCGCCTCCTGATCCAGGCGGTCGGCGCCCGCATCGACCAATATTACGACCGCGCCTGGGGCGCT
4792 4802 4812 4822 4832 4842 4852 4862
G E K R G T R L V V I G L H D M D E A A V R A A I
GGCGAAAAGCGCGGTACGCGCCTCGTCGTCATCGGCCTGCACGACATGGACGAGGCGGCGGTGCGCGCCGCGATC
4867 4877 4887 4897 4907 4917 4927 4937
T A L V * (SEQ. ID. NO: 47)
ACCGCGCTCGTGTAG (SEQ. ID. NO: 46)
4942 4952

FIG. 47G

COBW PROTEIN FIRST RESIDUE = 1
 LAST RESIDUE = 354

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	10	2.82	1470.68	3.86
2	LEU	L	32	9.04	3618.69	9.50
3	ILE	I	28	7.91	3166.35	8.31
4	MET	M	7	1.98	917.28	2.41
5	VAL	V	28	7.91	2773.92	7.28
6	SER	S	12	3.39	1044.38	2.74
7	PRO	P	11	3.11	1067.58	2.80
8	THR	T	21	5.93	2122.00	5.57
9	ALA	A	41	11.58	2912.52	7.64
10	TYR	Y	2	0.56	326.13	0.86
11	*	*	0	0.00	0.00	0.00
12	HIS	H	10	2.82	1370.59	3.60
13	GLN	Q	6	1.69	768.35	2.02
14	ASN	N	11	3.11	1254.47	3.29
15	LYS	K	15	4.24	1921.42	5.04
16	ASP	D	36	10.17	4140.97	10.87
17	GLU	E	27	7.63	3484.15	9.15
18	CYS	C	5	1.41	515.05	1.35
19	TRP	W	2	0.56	372.16	0.98
20	ARG	R	20	5.65	3122.02	8.19
21	GLY	G	30	8.47	1710.64	4.49
22	-	-	0	0.00	0.00	0.00

RESIDUES = 354
 MOLECULAR WEIGHT (MONOISOTOPIC) = 38097.4258
 MOLECULAR WEIGHT (AVERAGE) = 38121.1055
 INDEX OF POLARITY (%) = 44.63
 ISOELECTRIC POINT = 4.90
 OD 260 (1mg/ml) = 0.268 OD 280 (1mg/ml) = 0.354

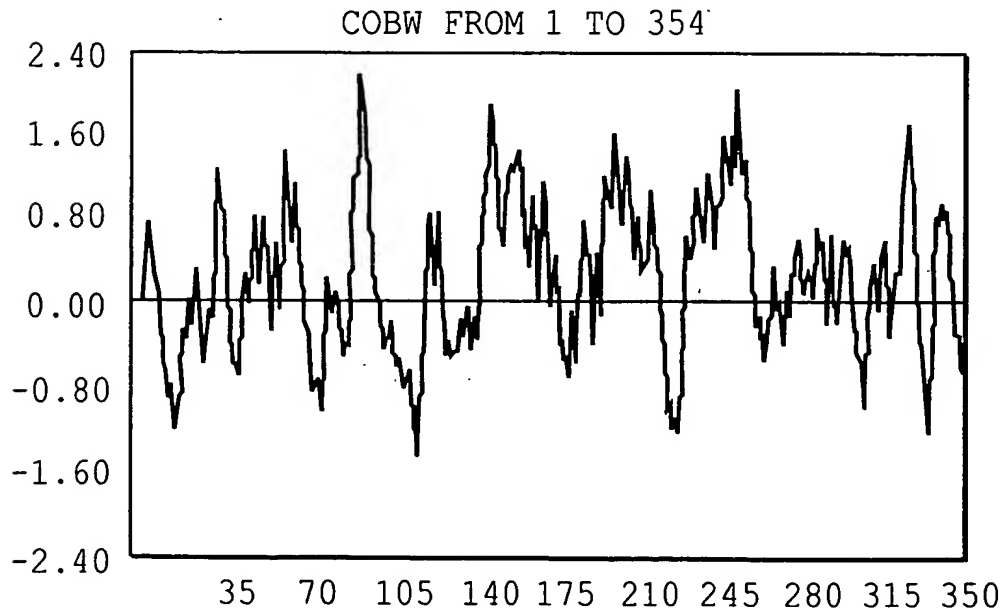


FIG. 47H

13144-BP SEQUENCE FROM 5060 TO 8887 cobN GENE

M H L L L A Q K G T I A D G N E A I D L G Q T P A
ATGCATCTGCTTCTCGCCCAGAAAGGAACGATCGCCGACGGCAACGAGGCGATCGACCTTGGGCAAACGCCGGCC
5060 5070 5080 5090 5100 5110 5120 5130
D I L F L S A A D T E L S S I A A A H G R R D G G
GATATCCTTTTCTATCGGCCGCCGACCCGAGCTCTCCTCGATCGCCGCGGCTCACGGCCGACGCGACGGAGGC
5135 5145 5155 5165 5175 5185 5195 5205
L S L R I A S L M S L M H P M S V D T Y V E R T A
TTGAGCCTGCGCATCGCCAGCCTGATGAGCCTGATGCACCCGATGTCGGTCGACACTTACGTCGAGCGCACGGCG
5210 5220 5230 5240 5250 5260 5270 5280
R H A K L I V V R P L G G A S Y F R Y L L E A L H
CGTCACGCCAAGCTGATCGTCGTCCGGCCGCTCGGTGGCGCCAGCTATTTCCGTTATCTGCTGGAAGCCCTGCAT
5285 5295 5305 5315 5325 5335 5345 5355
A A A V T H R F E I A V L P G D D K P D P G L E P
GCGGCTGCCGTCACCCATCGTTTCGAGATCGCGGTTCTGCCGGGTGACGACAAGCCGGATCCGGGGCTGGAGCCT
5360 5370 5380 5390 5400 5410 5420 5430
F S T V A A D D R Q R L W A Y F T E G G S D N A G
TTCTCCACCGTCGCAGCCGACGACCGCCAGCGCCTTTGGGCTTACTTCACCGAAGGCGGCTCGGACAATGCCGGG
5435 5445 5455 5465 5475 5485 5495 5505
L F L D Y A A A L V T G A E K P Q P A K P L L K A
CTGTTTCTCGACTATGCGGCCGCACTGGTCACAGGTGCGGAGAAGCCGAGCCGGCAAAGCCCCCTGTTGAAGGCC
5510 5520 5530 5540 5550 5560 5570 5580
G I W W P G A G V I G V S E W Q S L V Q G R M V A
GGCATCTGGTGGCCGGGTGCTGGTGTGATCGGCGTCAGCGAATGGCAGTCCCTTGTTTCAGGGACGGATGGTAGCG
5585 5595 5605 5615 5625 5635 5645 5655
R E G F E P P T V G I C F Y R A L V Q S G E T R P
AGGGAGGGATTCTGAACCCCCGACGGTCGGGATCTGCTTTTACCGCGGCTCGTGCAGAGTGGCGAGACACGGCCT
5660 5670 5680 5690 5700 5710 5720 5730
V E A L I D A L E A E G V R A L P V F V S S L K D
GTGGAGGCGCTGATCGATGCGCTGGAGGCTGAAGGTGTGCGGGCACTGCCGGTGTGTTGTCTCAAGCCTCAAGGAT
5735 5745 5755 5765 5775 5785 5795 5805
A V S V G T L Q A I F S E A A P D V V M N A T G F
GCCGTTTCCGTCGGCACGCTGCAGGCGATTTTTTCCGAGGCCGCACCCGACGTGGTGATGAACGCCACTGGCTTT
5810 5820 5830 5840 5850 5860 5870 5880
A V S S P G A D R Q P T V L E S T G A P V L Q V I
GCGGTCTCGTCGCCCCGGTGCCGACCGTCAGCCGACGGTCTGGAATCGACCGGTGCGCCGGTGTGTCAGGTGATT
5885 5895 5905 5915 5925 5935 5945 5955

FIG. 471

F S G S S R A Q W E T S P Q G L M A R D L A M N V
TTCTCCGGCTCGTCGCGGGCGCAATGGGAAACGTCGCCGAGGGGCTGATGGCGCGGACCTCGCCATGAACGTG
5960 5970 5980 5990 6000 6010 6020 6030
A L P E V D G R I L A R R V S F K A A S I Y D A K
GCACTCCCCGAAGTCGATGGCCGCATCCTTGCGCGCGCCGTCTCCTTCAAGGCGGCGTCGATCTATGACGCCAAG
6035 6045 6055 6065 6075 6085 6095 6105
V E A N I V G H E P L E G R V R F A A D L A V N W
GTGGAGGCCAATATCGTCGGCCATGAGCCGCTCGAAGGCCGGGTGCGCTTTGCCGCTGATCTTGCCGTCAACTGG
6110 6120 6130 6140 6150 6160 6170 6180
A N V R R A E P A E R R I A I V M A N Y P N R D G
GCGAACGTGCGCCGGGCGAGAGCCCGCCGAGCGCCGTATTGCCATCGTCATGGCCAACTATCCGAACCGCGACGGT
6185 6195 6205 6215 6225 6235 6245 6255
R L G N G V G L D T P A G T V E V L S A M A R E G
CGCCTCGGCAACGGTGTGCGGGCTCGACACGCCGCGCCGTACCGTCGAGGTGCTTAGCGCCATGGCGCGGGAAGGC
6260 6270 6280 6290 6300 6310 6320 6330
Y A V G E V P A D G D A L I R F L M A G P T N A A
TATGCGGTGCGTGAGGTTCGCCCGATGGCGACGCGCTGATCCGCTTTCTGATGGCCGGGCGACCAATGCGGCG
6335 6345 6355 6365 6375 6385 6395 6405
S H D R E I R E R I S L N D Y K T F F D S L P K Q
AGCCATGACCGTGAAATCCGCGAGCGTATTTCGCTGAACGATTACAAAACGTTCTTCGATTCGCTTCGGAAACAG
6410 6420 6430 6440 6450 6460 6470 6480
I K D E V A G R W G V P E A D P F F L D G A F A L
ATAAAGGATGAAGTTGCCGGTCGCTGGGGCGTGCCGGAGGCCGATCCCTTTTTCTCGATGGCGCCTTCGCGCTG
6485 6495 6505 6515 6525 6535 6545 6555
P L A R F G E V I V G I Q P A R G Y N I D P K E S
CCGCTCGCCCGCTTCGGCGAGGTGATCGTCGGCATCCAACCGCGCGCGGCTACAACATCGATCCGAAGGAAAGC
6560 6570 6580 6590 6600 6610 6620 6630
Y H S P D L V P P H G Y L A F Y A F L R Q Q F G A
TACCATTCCCCGGACCTCGTGCCGCCGATGGCTATCTCGCCTTCTACGCCTTCTGCGCCAGAGTTCCGAGCG
6635 6645 6655 6665 6675 6685 6695 6705
Q A I V H M G K H G N L E W L P G K A L A L S E T
CAGGCGATCGTCCACATGGGCAAGCACGGCAATCTCGAATGGTGCCGGGCAAGGCGCTGGCGCTGTCGGAAACC
6710 6720 6730 6740 6750 6760 6770 6780
C Y P E A I F G P L P H I Y P F I V N D P G E G T
TGCTATCCCGAAGCGATCTTCGGGCCGCTGCCGCACATCTATCCCTTCATCGTCAACGATCCGGGCGAAGGTACG
6785 6795 6805 6815 6825 6835 6845 6855

FIG. 47J

Q A K R R T S A V I I D H L T P P L T R A E S Y G
CAGGCCAAGCGCCGACCGAGCGCCGTCATCATCGACCACCTGACCCCGCCCTTGACGCGCGCCAATCCTACGGC
6860 6870 6880 6890 6900 6910 6920 6930
P L K D L E A L V D E Y Y D A A G G D P R R L R L
CCGCTCAAGGATCTGGAAGCGCTCGTCGACGAATATTACGACGCCGCGGCGGTGATCCGCGCCGCTCAGGCTG
6935 6945 6955 6965 6975 6985 6995 7005
L S R Q I L D L V R D I G L D S D A G I D R G D S
CTCAGCCGCCAGATCCTCGATCTCGTGC GCGACATCGGCCTCGACAGCGACGCAGGCATCGACAGGGGCGACAGC
7010 7020 7030 7040 7050 7060 7070 7080
D D K A L E K L D A Y L C D L K E M Q I R D G L H
GACGACAAGGCGCTGAAAAGCTCGACGCCTATCTCTGCGACCTCAAGGAAATGCAGATCCGCGACGGCCTGCAC
7085 7095 7105 7115 7125 7135 7145 7155
I F G V A P E G R L L T D L T V A L A R V P R G L
ATCTTCGGCGTTGCGCCGAAGGGCGGTTGTTGACGGACCTCACCGTAGCGCTGGCGCGCTGCCCCGAGGTCTC
7160 7170 7180 7190 7200 7210 7220 7230
G E G G D Q S L Q R A I A A D A G L R G F A I P T
GGCGAGGGCGGCGACCGAGCCTGCAGCGGGCGATCGCAGCGGATGCGGGGCTGCGTGGGTTTGCTATTCCCACC
7235 7245 7255 7265 7275 7285 7295 7305
S A G G N P A R D A Q P F D P L D C V M S D T W T
TCGGCGGGGGGCAACCCCGCACGCGACGCCCAACCCTTCGACCCGCTCGACTGCGTCATGTCCGACACCTGGACA
7310 7320 7330 7340 7350 7360 7370 7380
G P K P S I L A D L S D A P W R T A G D T V E R I
GGCCCCAAACCGTCCATCCTCGCTGACCTCTCGGACGCCCCCTGGCGCACCGCCGGCGATACGGTCGAGCGCATC
7385 7395 7405 7415 7425 7435 7445 7455
E L L A A N L V S G E L A C P D H W A N T R A V L
GAGTTGCTTGCCGCAAATCTCGTGTGCGGTGAAGTGGCTTGCCCGGACCACTGGGCCAACACCCGCGCCGTGCTC
7460 7470 7480 7490 7500 7510 7520 7530
G E I E T R L K P S I S N S G A A E M T G F L T G
GGCGAAATCGAAACGCGCCTGAAGCCGTCGATTTCAAACTCGGGTGCCGCGGAGATGACCGGCTTCCTACCGGT
7535 7545 7555 7565 7575 7585 7595 7605
L S G R F V A P G P S G A P T R G R P D V L P T G
CTCAGCGGCCGCTTCGTGCCCCCGGTCCATCGGGCGCGCCGACGCGCGCCGCGCGGATGTGTTGCCGACGGGG
7610 7620 7630 7640 7650 7660 7670 7680
R N F Y S V D S R A V P T P A A Y E L G K K S A E
CGCAATTTCTACTCGGTCGACAGCCGCGCCGTGCCGACGCCGGCGGCTTACGAGCTTGGAAGAAATCGGCCGAG
7685 7695 7705 7715 7725 7735 7745 7755

FIG. 47K

L L I R R Y L Q D H G E W P S S F G L T A W G T A
CTTCTGATCCGCCGCTACCTGCAGGACCATGGCGAATGGCCGTCCTCCTTTGGCCTGACCGCCTGGGGCACGGCG
7760 7770 7780 7790 7800 7810 7820 7830
N M R T G G D D I A Q A L A L I G A K P T W D M V
AACATGCGCACCGCGCGGACGACATCGCCCAGGCCCTGGCGCTGATCGGCGCCAAGCCCACCTGGGACATGGTC
7835 7845 7855 7865 7875 7885 7895 7905
S R R V M G Y E I V P L A V L G R P R V D V T L R
TCTCGCCGGGTGATGGGCTACGAGATCGTGCCGCTCGCAGTCCTCGCCGCCACGCGTCGACGTGACCTTGCGC
7910 7920 7930 7940 7950 7960 7970 7980
I S G F F R D A F P D Q I A L F D K A I R A V A L
ATTTCCGGCTTCTTCGCGATGCCTTCCCGGACCAGATCGCGCTCTTCGACAAGGCGATCCGCGCCGTCGCGCTG
7985 7995 8005 8015 8025 8035 8045 8055
E E D D A D N M I A A R M R A E S R R L E A E G V
GAGGAAGACGATGCCGACAACATGATCGCCGCACGCATCGGGGCGAAAGCCGGCGGCTGGAGGCCAAGGCGTG
8060 8070 8080 8090 8100 8110 8120 8130
E A A E A A R R A S Y R V F G R K P G A Y G A A L
GAAGCCGCCGAGGCCGCGCTCGCGCTCTACCGCGTCTTTGGCGCAAAGCCCGGTGCCTATGGCGCCGCCCTG
8135 8145 8155 8165 8175 8185 8195 8205
Q A L I D E K G W E T K A D L A E A Y L T W G A Y
CAGGCGCTGATCGACGAGAAGGGCTGGGAAACCAAAGCAGATCTCGCCGAGGCCTATCTTACCTGGGGCGCCTAT
8210 8220 8230 8240 8250 8260 8270 8280
A Y G A G E E G K A E R D L F E E R L R T I E A V
GCCTATGGCGCCGGCGAGGAGGGCAAGGCCGAGCGGATCTTTTCGAGGAGCGCCTGCGCACGATAGAGGCGGTG
8285 8295 8305 8315 8325 8335 8345 8355
V Q N Q D N R E H D L L D S D D Y Y Q F E G G M S
GTGCAGAACCAGGACAACCGCGAGCACGATCTGCTCGACAGCGACGACTACTACAGTTCGAAGGCGGCATGAGC
8360 8370 8380 8390 8400 8410 8420 8430
A A A E Q L G G H R P A I Y H N D H S R P E K P V
GCTGCCGCCGAACAGCTCGGCGGTCACCGTCCGGCGATCTACCACAACGACCATTCCCGTCCGGAAAAGCCTGTG
8435 8445 8455 8465 8475 8485 8495 8505
I R S L E E E I G R V V R A R V V N P K W I D G V
ATCCGGTCGCTCGAAGAAGAGATCGGCCGCGTGGTCCGGGCCCCGCTCGTCAATCCCAAGTGGATCGATGGCGTC
8510 8520 8530 8540 8550 8560 8570 8580
M R H G Y K G A F E I A A T V D Y M F A F A A T T
ATGCCACCGGATACAAGGGCGCCTTCGAGATCGCTGCCACGGTCGACTACATGTTGCCTTTGCCGCGACCACG
8585 8595 8605 8615 8625 8635 8645 8655

FIG. 47L

G A V R D H H F E A A Y Q A F I V D E R V A D F M
GGTGC GGTGCGCGACCATCATTTTCGAGGCCGCTTATCAGGCGTTCATTGTCGACGAGCGCGTGGCTGACTTCATG
8660 8670 8680 8690 8700 8710 8720 8730
R D K N P A A F A E L A E R L L E A I D R N L W T
CGCGACAAGAACCCGGCCGCTTTGCCGAGCTTGCCGAACGCCTGCTTGAAGCAATCGACCGCAATCTCTGGACG
8735 8745 8755 8765 8775 8785 8795 8805
P R S N S A R F E L A G I G T A A T R L R A G N E
CCGCGCTCGAATTCGGCGCGGTTTGAACCTGCCGGCATCGGCACGGCAGCAACCCGGCTTCGTGCCGGCAATGAA
8810 8820 8830 8840 8850 8860 8870 8880
* (SEQ ID NO: 49)
TAG (SEQ ID NO: 48)
8885

FIG. 47M

COBN PROTEIN FIRST RESIDUE = 1
 LAST RESIDUE = 1275

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	48	3.76	7059.28	5.12
2	LEU	L	121	9.49	13683.17	9.92
3	ILE	I	60	4.71	6785.04	4.92
4	MET	M	24	1.88	3144.97	2.28
5	VAL	V	82	6.43	8123.61	5.89
6	SER	S	64	5.02	5570.05	4.04
7	PRO	P	76	5.96	7376.01	5.35
8	THR	T	53	4.16	5355.53	3.88
9	ALA	A	180	14.12	12786.68	9.27
10	TYR	Y	35	2.75	5707.22	4.14
11	*	*	0	0.00	0.00	0.00
12	HIS	H	24	1.88	3289.41	2.38
13	GLN	Q	32	2.51	4097.87	2.97
14	ASN	N	30	2.35	3421.29	2.48
15	LYS	K	34	2.67	4355.23	3.16
16	ASP	D	90	7.06	10352.42	7.50
17	GLU	E	85	6.67	10968.62	7.95
18	CYS	C	5	0.39	515.05	0.37
19	TRP	W	18	1.41	3349.43	2.43
20	ARG	R	99	7.76	15454.01	11.20
21	GLY	G	115	9.02	6557.47	4.75
22	-	-	0	0.00	0.00	0.00

RESIDUES = 1275
 MOLECULAR WEIGHT (MONOISOTOPIC) = 137970.5000
 MOLECULAR WEIGHT (AVERAGE) = 138055.8594
 INDEX OF POLARITY (%) = 40.08
 ISOELECTRIC POINT = 5.42
 OD 260 (1mg/ml) = 0.693 OD 280 (1mg/ml) = 1.027

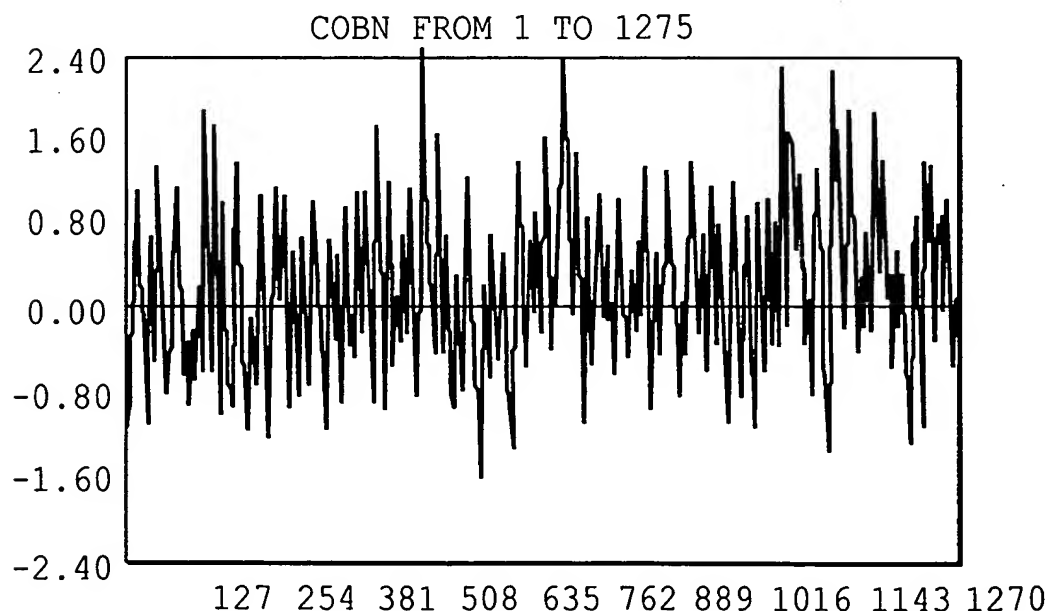


FIG. 47N

13144-BP SEQUENCE 9034 TO 9678 cobO GENE

```

      M S D E T T V G G E A P A E K D D A R H A M K M A
      ATGAGCGACGAGACGACAGTAGGCGGCGAAGCCCCGCCGAGAAGGACGATGCCCCCACGCCATGAAGATGGCG
9034      9044      9054      9064      9074      9084      9094      9104
      K K K A A R E K I M A T K T D E K G L I I V N T G
      AAGAAGAAGGCAGCACGCGAAAAGATCATGGCGACGAAGACCGACGAGAAGGGTCTGATCATCGTCAACACCGGC
9109      9119      9129      9139      9149      9159      9169      9179
      K G K G K S T A G F G M I F R H I A H G M P C A V
      AAAGGCAAGGGCAAGTCGACCGCCGGCTTCGGCATGATCTTCGCCATATCGCCACGGCATGCCCTGCGCCGTC
9184      9194      9204      9214      9224      9234      9244      9254
      V Q F I K G A M A T G E R E L I E K H F G D V C Q
      GTGCAGTTCATCAAGGGTTCGATGGCAACCGGCGAGCGCGAGTTGATCGAGAAGCATTTCGGCGATGTCTGCCAG
9259      9269      9279      9289      9299      9309      9319      9329
      F Y T L G E G F T W E T Q D R A R D V A M A E K A
      TTCTACACGCTCGGCGAGGGCTTCACCTGGGAAACGCAGGATCGCGCCCGCGATGTTGCGATGGCTGAAAAGGCC
9334      9344      9354      9364      9374      9384      9394      9404
      W E K A K E L I R D E R N S M V L L D E I N I R L
      TGGGAGAAGGCGAAGGAAGTATCCGTGACGAGCGCAACTCGATGGTGCTGCTCGACGAGATCAACATTGCTCTG
9409      9419      9429      9439      9449      9459      9469      9479
      R Y D Y I D V A E V V R F L K E E K P H M T H V V
      CGCTACGACTACATCGACGTCGCCGAAGTGGTGCGCTTCCTGAAGGAAGAAAAGCCGCACATGACGCATGTGGTG
9484      9494      9504      9514      9524      9534      9544      9554
      L T G R N A K E D L I E V A D L V T E M E L I K H
      CTCACCGGCCGCAACGCGAAAGAAGACCTGATCGAAGTCGCCGATCTCGTCACTGAGATGGAGCTGATCAAGCAT
9559      9569      9579      9589      9599      9609      9619      9629
      P F R S G I K A Q Q G V E F *      (SEQ ID NO: 51)
      CCGTTCGGTTCGGCATCAAGGCGCAGAGGGCGTGGAGTTCTGA      (SEQ ID NO: 50)
9634      9644      9654      9664      9674
```

FIG. 470

COBO PROTEIN			FIRST RESIDUE = 1		LAST RESIDUE = 214	
			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	9	4.21	1323.62	5.51
2	LEU	L	12	5.61	1357.01	5.65
3	ILE	I	14	6.54	1583.18	6.59
4	MET	M	11	5.14	1441.45	6.00
5	VAL	V	15	7.01	1486.03	6.19
6	SER	S	4	1.87	348.13	1.45
7	PRO	P	4	1.87	388.21	1.62
8	THR	T	13	6.07	1313.62	5.47
9	ALA	A	23	10.75	1633.85	6.80
10	TYR	Y	3	1.40	489.19	2.04
11	*	*	0	0.00	0.00	0.00
12	HIS	H	7	3.27	959.41	4.00
13	GLN	Q	5	2.34	640.29	2.67
14	ASN	N	4	1.87	456.17	1.90
15	LYS	K	21	9.81	2689.99	11.20
16	ASP	D	13	6.07	1495.35	6.23
17	GLU	E	23	10.75	2967.98	12.36
18	CYS	C	2	0.93	206.02	0.86
19	TRP	W	2	0.93	372.16	1.55
20	ARG	R	12	5.61	1873.21	7.80
21	GLY	G	17	7.94	969.36	4.04
22	-	-	0	0.00	0.00	0.00

RESIDUES = 214
 MOLECULAR WEIGHT (MONOISOTOPIC) = 24012.2500
 MOLECULAR WEIGHT (AVERAGE) = 24027.6973
 INDEX OF POLARITY (%) = 47.66
 ISOELECTRIC POINT = 6.94
 OD 260 (1mg/ml) = 0.443 OD 280 (1mg/ml) = 0.612

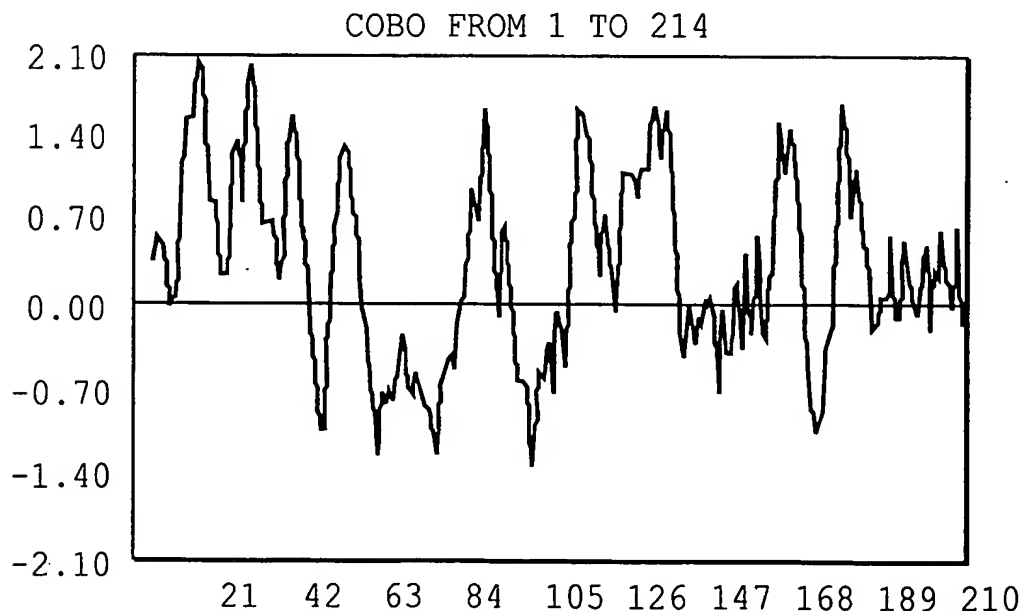


FIG. 47P

NH₂-TERMINAL SEQUENCE OF SUMT OF M. ivanovii

VVYLVGAGPGDPELITLKAVNVLK-ADVVL

(AMINO ACID FRAGMENT 2-31

923 946

OF SEQ ID NO: 54)

SENSE OLIGONUCLEOTIDE 946 (27-mer)

		P	G	D	P	E	L		(AMINO ACIDS 10-15 OF SEQ ID NO: 54)
5'	CGCGGAATTC	CCA	GGA	GAT	CCA	GAA	CT	3'	(SEQ ID NO: 56)
	EcoRI	T	T	C	T	G			
		C	C		C				
		G	G		G				

SENSE OLIGONUCLEOTIDE 923 (27-mer)

		V	Y	L	V	G	A		(AMINO ACIDS 3-8 OF SEQ ID NO: 54)
5'	CGCGGAATTC	GTA	TAT	CTA	GTA	GGA	GC	3'	(SEQ ID NO: 57)
	EcoRI	G	C	T	T	T			
					G				

NH₂-TERMINAL SEQUENCE OF A FRAGMENT DERIVED FROM A
 TRYPTIC DIGESTION OF M. ivanovii SUMT

IITGTLENIAGK

(AMINO ACIDS 201-212 OF

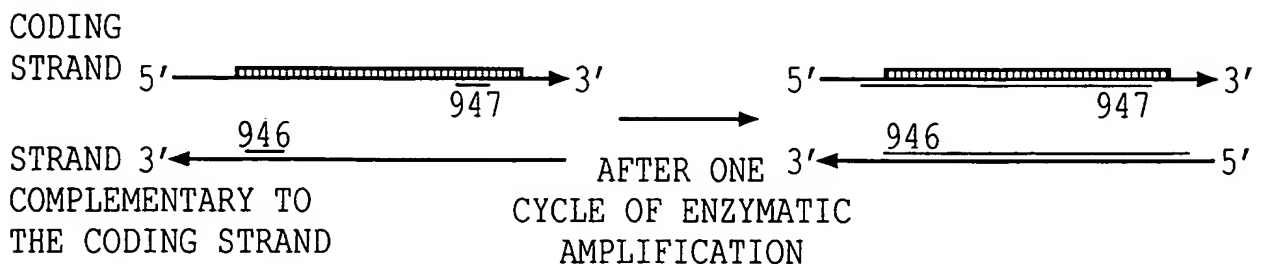
947

SEQ ID NO: 54)

ANTISENSE OLIGONUCLEOTIDE 947 (25-mer)

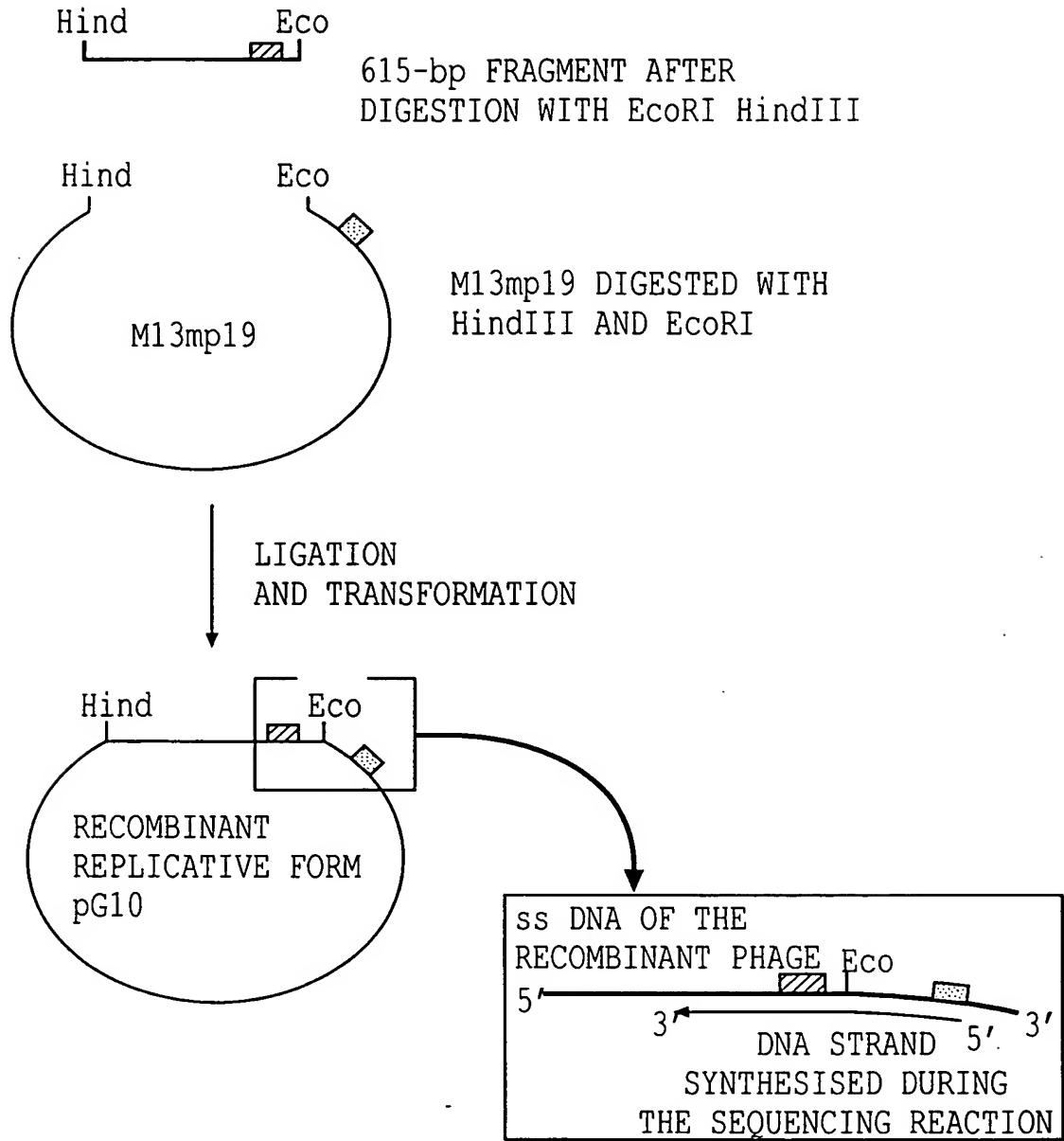
		N	E	L	T	G		(CO ₂ →NH ₂ ORIENTATION OF AMINO
5'	CGCGAAGCTT	GTT	TTC	TAG	AGT	ACC	3'	ACIDS 207-203 OF SEQ ID NO: 54)
	HindIII	A	C	A	T	T		(SEQ ID NO: 58)
					G	C		
					G			

FIG. 48A



M. ivanovii SUMT STRUCTURAL GENE (SEQ ID NO: 54)

FIG. 48B



- HYBRIDISATION SITE OF THE PRIMER -20 OF THE ss DNA OF PHAGE M13mp19
- ▨ SEQUENCE COMPLEMENTARY TO THE SENSE OLIGONUCLEOTIDE 946

FIG. 49

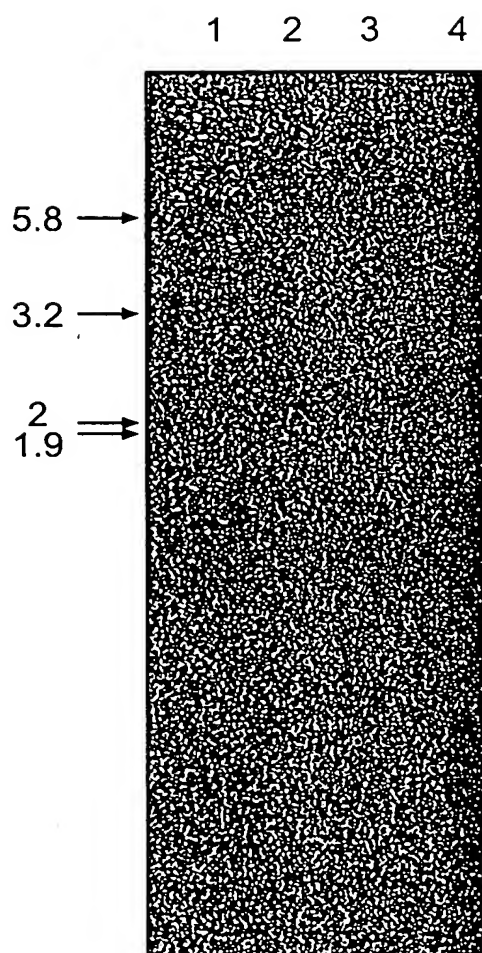


FIG. 50

10	20	30	40	50	60
CCATAATTCT	TTTATAATTT	AAACGGTGAA	CACATGGTAG	TTTATTTAGT	AGGTGCGGGT
GGTATTAAGA	AAATATTAAA	TTTGCCACTT	GTGTACCATC	AAATAAATCA	TCCACGCCCA
70	80	90	100	110	120
CCAGGAGATC	CCGAACCTAT	CACTCTCAAA	GCTGTAAACG	TGTTAAAAAA	AGCGGATGTT
GGTCCTCTAG	GGCTTGAATA	GTGAGAGTTT	CGACATTTCG	ACAATTTTTT	TCGCCTACAA
130	140	150	160	170	180
GTACTGTACG	ACAAACCTGC	AAATGAAGAA	ATTTTAAAGT	ATGCTGAAGG	TGCAAAACTA
CATGACATGC	TGTTTGGACG	TTTACTTCTT	TAAAATTTC	TACGACTTCC	ACGTTTTGAT
190	200	210	220	230	240
ATATATGTCG	GAAAACAAGC	AGGACATCAT	TACAAATCTC	AAAATGAAAT	CAATACTCTT
TATATACAGC	CTTTTGTTTCG	TCCTGTAGTA	ATGTTTAGAG	TTTTACTTTA	GTTATGAGAA
250	260	270	280	290	300
CTTGTTGAAG	AAGCAAAAGA	AAATGATTTA	GTAGTACGCC	TTAAAGGTGG	AGACCCCTTT
GAACAACTTC	TTCGTTTTCT	TTTACTAAAT	CATCATGCGG	AATTTCCACC	TCTGGGGAAA
310	320	330	340	350	360
GTATTTGGAA	GAGGAGGCGA	GGAAATTCTG	GCCCTTGTAG	AAGAAGGAAT	TGATTTTGAG
CATAAACCTT	CTCCTCCGCT	CCTTTAAGAC	CGGGAACATC	TTCTTCCTTA	ACTAAAACTC
370	380	390	400	410	420
TTAGTTCCAG	GGGTAACCTC	TGCAATTGGA	GTTCCAACAA	CAATTGGGCT	TCCAGTTACT
AATCAAGGTC	CCCATTGAAG	ACGTTAACCT	CAAGGTTGTT	GTTAACCCGA	AGGTCAATGA
430	440	450	460	470	480
CATAGAGGTG	TTGCAACATC	GTTTACAGTT	GTTACAGGTC	ATGAAGACCC	AACAAAATGC
GTATCTCCAC	AACGTTGTAG	CAAATGTCAA	CAATGTCCAG	TACTTCTGGG	TTGTTTTACG
490	500	510	520	530	540
AAGAAACAGG	TAGGATGGGA	CTTTAAAGCA	GATACTATTG	TAATACTTAT	GGGTATTGGA
TTCTTTGTCC	ATCCTACCCT	GAAATTTCTG	CTATGATAAC	ATTATGAATA	CCCATAACCT
550	560	570	580	590	600
AATTTAGCTG	AAAATACAGC	AGAAATTATG	AAACATAAAG	ATCCTGAAAC	TCCAGTTTGT
TTAAATCGAC	TTTTATGTCG	TCTTTAATAC	TTTGTATTTC	TAGGACTTTG	AGGTCAAACA

FIG. 51A

610	620	630	640	650	660
GTAATTGAAA	ATGGTACGAT	GGAAGGTCAA	AGGATAATAA	CGGGCACACT	GGAAAATATA
CATTAAC TTT	TACCATGCTA	CCTTCCAGTT	TCCTATTATT	GCCCGTGTGA	CCTTTTATAT
670	680	690	700	710	720
GCTGGAAAGG	ATATTAAACC	TCCTGCTTTA	GTGGTATTGG	AAATGTTGTC	AATGTTTTTA
CGACCTTTCC	TATAATTTGG	AGGACGAAAT	CACCATAACC	TTTACAACAG	TTACAAAAAT
730	740	750	760	770	780
AAGAAATGAA	TCAAATCAGT	GGCTGATCTA	TTAAGAAGGC	AATATCATGA	ATGGATTAGA
TTCTTTACTT	AGTTTAGTCA	CCGACTAGAT	AATTCTTCCG	TTATAGTACT	TACCTAATCT
790	800	810	820	830	840
AGGTAAAAAA	ATTGTTATAA	CAAGACCTGC	TGAAAGGGCT	AAAGACTCAG	TTGAAATGGT
TCCATTTTTT	TAACAATATT	GTTCTGGACG	ACTTTCCCGA	TTTCTGAGTC	AACTTTACCA
850	860	870	880	890	900
AAAATCTTAT	GGAGCAGTTC	CAATTGTAAC	TCCTACAATT	GAACTCAAAG	ATTCCAAGCC
TTTTAGAATA	CCTCGTCAAG	GTTAACATTG	AGGATGT TAA	CTTGAGTTTC	TAAGGTTCGG
910	920	930	940	950	960
AGAAGAAGTG	ATAAAATTAT	GTAATATGAT	AAATGAACCT	TGATTGGCCT	TATAT
TCTTCTTCAC	TATTTTAATA	CATTATACTA	TTTACTTGGA	ACTAACCGGA	ATATA

FIG. 51B

corA GENE AND CORA PROTEIN (SUMT) OF METHANOBACTERIUM
IVANOVII SEQUENCE OF 955-BP FRAGMENT
FROM 34 TO 729

```

M V V Y L V G A G P G D P E L I T L K A V N V L K
ATGGTAGTTTATTTAGTAGGTGCGGGTCCAGGAGATCCCGAACTTATCACTCTCAAAGCTGTAAACGTGTTAAAA
34      44      54      64      74      84      94      104
K A D V V L Y D K P A N E E I L K Y A E G A K L I
AAAGCGGATGTTGTACTGTACGACAAACCTGCAAATGAAGAAATTTTAAAGTATGCTGAAGGTGCAAACTAATA
109     119     129     139     149     159     169     179
Y V G K Q A G H H Y K S Q N E I N T L L V E E A K
TATGTCGGAACAAGCAGGACATCATTACAAATCTCAAAATGAAATCAATACTCTTCTTGTGAAGAAGCAAAA
184     194     204     214     224     234     244     254
E N D L V V R L K G G D P F V F G R G G E E I L A
GAAAATGATTTAGTAGTACGCCTTAAAGGTGGAGACCCCTTTGTATTTGGAAGAGGAGCGAGGAAATCTGGCC
259     269     279     289     299     309     319     329
L V E E G I D F E L V P G V T S A I G V P T T I G
CTTGTAGAAGAAGGAATTGATTTTGAGTTAGTTCCAGGGGTAACCTCTGCAATTGGAGTTCCAACAACAATTGGG
334     344     354     364     374     384     394     404
L P V T H R G V A T S F T V V T G H E D P T K C K
CTTCCAGTTACTCATAGAGGTGTTGCAACATCGTTTACAGTTGTTACAGGTCATGAAGACCCAACAAAATGCAAG
409     419     429     439     449     459     469     479
K Q V G W D F K A D T I V I L M G I G N L A E N T
AAACAGGTAGGATGGGACTTTAAAGCAGATACTATTGTAATACTTATGGGTATTGGAAATTTAGCTGAAAATACA
484     494     504     514     524     534     544     554
A E I M K H K D P E T P V C V I E N G T M E G Q R
GCAGAAATTATGAAACATAAAGATCCTGAAACTCCAGTTTGTGTAATTGAAATGGTACGATGGAAGGTCAAAGG
559     569     579     589     599     609     619     629
I I T G T L E N I A G K D I K P P A L V V L E M L
ATAATAACGGGCACACTGGAAAATATAGCTGGAAAGGATATTAAACCTCCTGCTTTAGTGGTATTGGAAATGTTG
634     644     654     664     674     684     694     704
S M F L K K *
TCAATGTTTTTAAAGAAATGA
709     719     729
```

FIG. 52A

CORA PROTEIN

			NUMBER	NO. %	WEIGHT	WEIGHT %
1	PHE	F	6	2.60	882.41	3.54
2	LEU	L	22	9.52	2487.85	9.99
3	ILE	I	17	7.36	1922.43	7.72
4	MET	M	6	2.60	786.24	3.16
5	VAL	V	26	11.26	2575.78	10.34
6	SER	S	4	1.73	348.13	1.40
7	PRO	P	12	5.19	1164.63	4.68
8	THR	T	16	6.93	1616.76	6.49
9	ALA	A	16	6.93	1136.59	4.56
10	TYR	Y	5	2.16	815.32	3.27
11	*	*	0	0.00	0.00	0.00
12	HIS	H	5	2.16	685.29	2.75
13	GLN	Q	4	1.73	512.23	2.06
14	ASN	N	9	3.90	1026.39	4.12
15	LYS	K	20	8.66	2561.90	10.29
16	ASP	D	11	4.76	1265.30	5.08
17	GLU	E	21	9.09	2709.89	10.88
18	CYS	C	2	0.87	206.02	0.83
19	TRP	W	1	0.43	186.08	0.75
20	ARG	R	4	1.73	624.40	2.51
21	GLY	G	24	10.39	1368.52	5.50
22	-	-	0	0.00	0.00	0.00

RESIDUES = 231
 MOLECULAR WEIGHT (MONOISOTOPIC) = 24900.1855
 MOLECULAR WEIGHT (AVERAGE) = 24915.9766
 INDEX OF POLARITY (%) = 40.69
 ISOELECTRIC POINT = 5.45
 OD 260 (1mg/ml) = 0.304 OD 280 (1mg/ml) = 0.462

FIG. 52B

CORA FROM FIRST RESIDUE = 1
LAST RESIDUE = 231

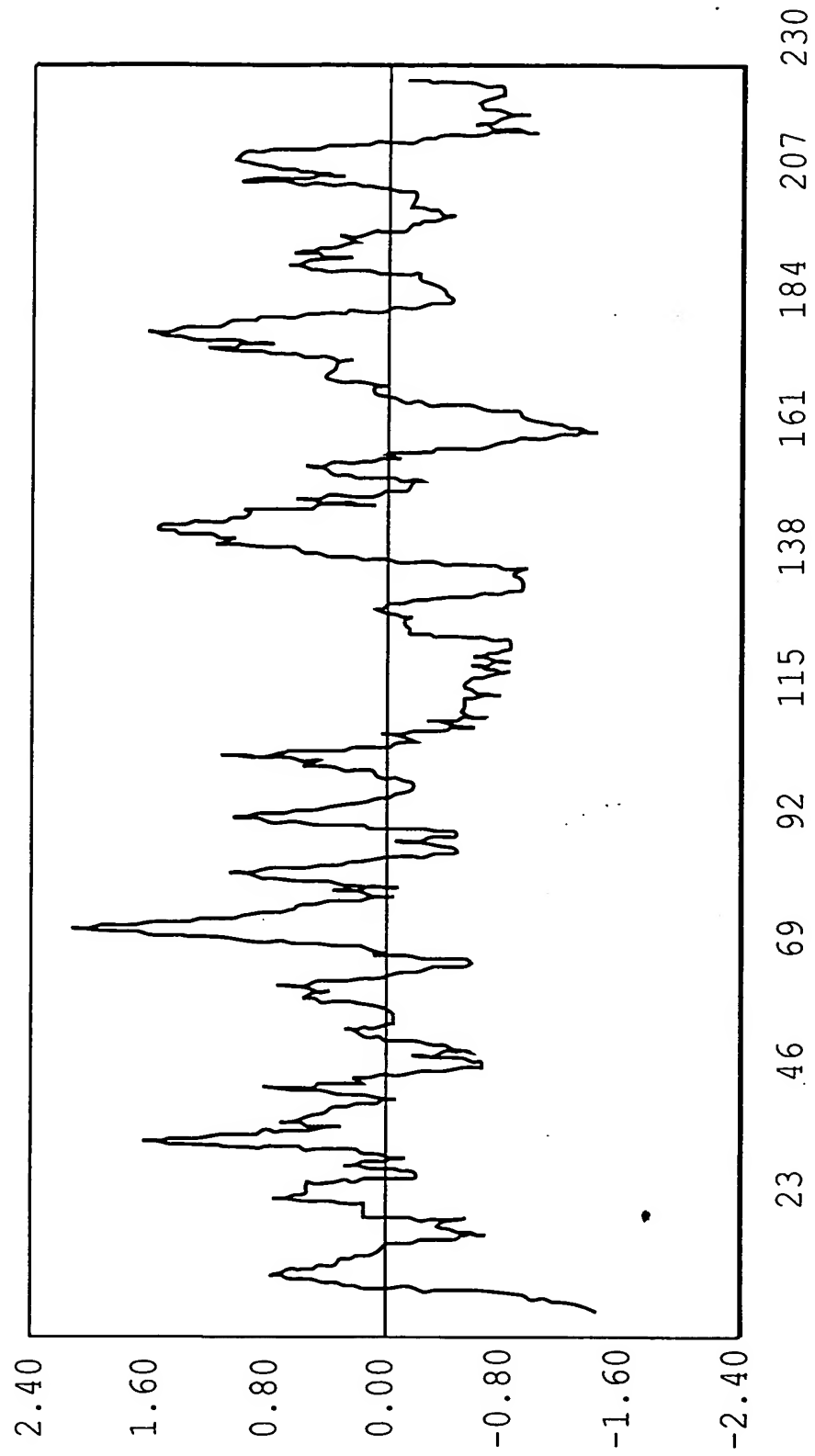


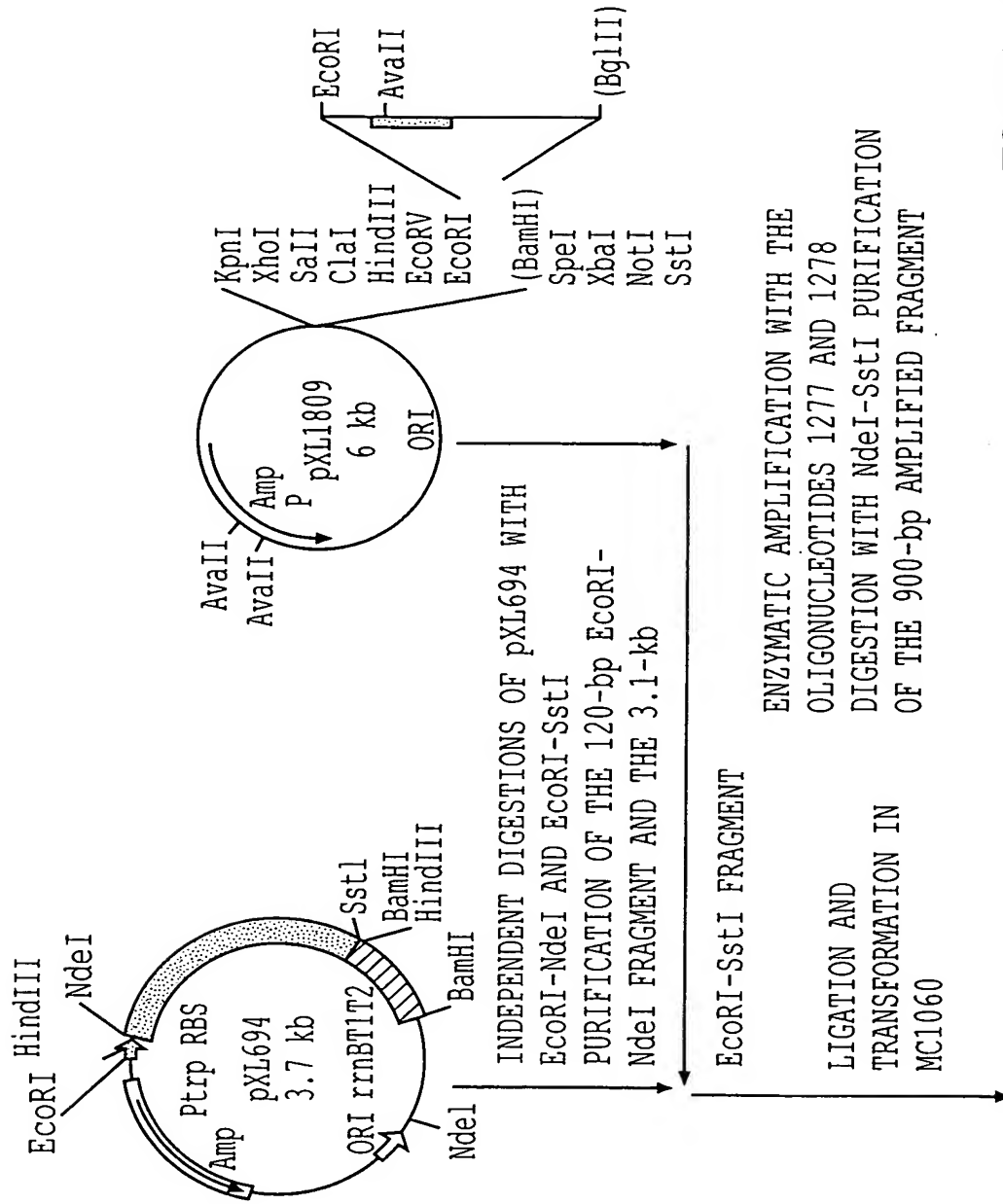
FIG. 52C

M. IVANOVII CORA	3 TO 227
P. DENITRIFICANS COBA	17 TO 251

10	20	30	40	50	60	70
VYLVGAGPGDPELITLKAVNVLKKADVLYDKPANEEILKYAE-GAKLIYVGKQAGHHYKSQNEINTLLV						
=====	=====	=====	=====	=====	=====	=====
VWLVGAGPGDPGLLTLHAANALRQADVIVHDAVNEDCLKLARPGAVLEFAGKRGKPSPKQORDISLRLV						
20	30	40	50	60	70	80
80	90	100	110	120	130	140
EEAKENDLVVRLKGGDPFVFGRGEEELALVEEGIDFELVPGVTSAGVPTTIGLPVTHRGVATSFTVVT						
==	==	==	==	==	==	==
ELARAGNRVRLKGGDPFVFGRGEEALTVEHQVPPRIVPGITAGIGGLAYAGIPVTHREVNHAVTFLT						
90	100	110	120	130	140	150
150	160	170	180	190	200	
GHDPTKCKKQVGWD-FKADT-IVIL-MGIGNLAENTAIEIMKH-KDPETPVCVIENTMEGQRIITGTL-						
==	==	==	==	==	==	==
GHDSSGLVPDRINWQGIASGSPVIVMYMAMKHIGAITANLIAGGRSPDEPVAFVCNAATPQQAVLETTLA						
160	170	180	190	200	210	220
210	220					
--E-NIAGKDIKPPALVVL-EMLSM						
=	=	=	=	=	=	=
RAEADVAAGLEPPAIVVGEVVRL						
230	240	250				

PERCENTAGE OF STRICT HOMOLOGY : 40.4

FIG. 53



TO FIG. 54B

FIG. 54A

TO FIG. 54B



PURIFICATION OF PSEUDOMONAS DENITRIFICANS COBINAMIDE KINASE - COBINAMIDEPHOSPHATE
 GUANYLYL TRANSFERASE

PURIFICATION STEP	VOL (ml)	PROTEIN (mg)	COBINAMIDE KINASE			COBINAMIDEPHOSPHATE GUANYLYL TRANSFERASE	RATIO OF sp. act. 2/1
			sp. act. 1 (U/mg of protein)	YIELD (%)	PURIFICATION FACTOR		
CRUDE EXTRACT*	4.5	120	16	-	-	214	13
MONO Q HR 10/10 ELUATE	9.0	8.98	188	88	12	-	-
HYDROXYAPATITE ELUATE	2.0	4.55	325	77	20	3640	11
PHENYL-SUPEROSE ELUATE	2.0	1.51	560	44	35	-	-
MONO Q HR 5/5 ELUATE	3.0	0.90	786	37	49	11282	14

*FROM 1 G OF WET SC510 PXL622 CELLS CULTURED ON PS4 MEDIUM (CAMERON ET AL., 1989) WITHOUT COBAL T.

FIG. 55